University of Dundee

DOCTOR OF PHILOSOPHY

Development and validation of an inventory (Dundee Barometer of Institutional Professionalism) to measure the professionalism culture of medical schools in the UK

Chandratilake, Madawa Nilupathi

Award date:
2013

Awarding institution:
University of Dundee

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 17. Feb. 2017

DOCTOR OF MEDICINE

Development and validation of an inventory (Dundee Barometer of Institutional Professionalism) to measure the professionalism culture of medical schools in the UK

Madawa Nilupathi Chandratilake

2013

University of Dundee

Conditions for Use and Duplication
Copyright of this work belongs to the author unless otherwise identified in the body of the thesis. It is permitted to use and duplicate this work only for personal and non-commercial research, study or criticism/review. You must obtain prior written consent from the author for any other use. Any quotation from this thesis must be acknowledged using the normal academic conventions. It is not permitted to supply the whole or part of this thesis to any other person or to post the same on any website or other online location without the prior written consent of the author. Contact the Discovery team (discovery@dundee.ac.uk) with any queries about the use or acknowledgement of this work.

Development and Validation of an Inventory (Dundee Barometer of Institutional Professionalism) to Measure the Professionalism Culture of Medical Schools in the UK

Madawa Nilupathi Chandratilake

A Thesis Submitted for the Degree of Doctor of Philosophy

University of Dundee

October 2012

Declaration

I hereby declare that I am the author of this thesis; that the work of which this thesis is a record has been done by myself, and that it has not previously been accepted for a higher degree.

Madawa N. Chandratilake

October 2012

Acknowledgements

I extend my deep and sincere gratitude to

Dr Sean McAleer and Dr John Gibson, my supervisors: you were always available and accessible. Both of you provided the intellectual challenge, wisdom, motivation, guidance and feedback to make this project a success. I very much enjoyed working with you.

Prof Margery Davis, former director of the Centre for Medical Education, Dundee: you offered me the doctoral project which I consider as a life-time opportunity, and guided me throughout to completion.

Mrs Sue Roff, my former colleague: you were always there when I needed guidance, feedback and support.

Prof Charlotte Rees, Director of the Centre for Medical Education, Dundee: you always encouraged me to achieve multiple goals with this doctoral project by facilitating opportunities to present at conferences and workshops, and by peer reviewing my manuscripts.

Prof Nilanthi de Silva, Prof Rajitha Wickermasinghe and **Prof Janaka de Silva,** my superiors at the Faculty of Medicine, University of Kelaniya, Sri Lanka: you worked untiringly to make the necessary arrangements from the University of Kelaniya to complete my doctoral project smoothly and without any interruption.

Prof David Wall, external tutor at the Centre for Medical Education, Dundee: your advice, guidance and assistance made the use of generalisability theory a lot clearer.

The UK general public, healthcare professionals, colleagues and students of the Dundee Medical School, the participants in the project: Without your help and support, this project would not have been a success.

Academic and non-academic colleagues at the Centre for Medical Education, Dundee: you supported me in every possible way to achieve my goal of completing this project within the time limit.

Mrs Dilrukshi Chandratilake, my wife, and **Master Thisath Chandratilake**, my son: you both were wonderful. You made no complaints about my long working hours or missing family time. But you always shared your love and happiness in every little opportunity we had.

Mrs Dhammika Chandratilake, my mother: you were my first teacher. The discipline and commitment you imparted were really tested during this project.

Dedication

To my beautiful wife, *Dilu* and adorable son, *Thisath*

Preface

The thesis is presented in four sections (A-D) to enhance clarity and ease of understanding. Each section comprises several chapters, which have been divided into appropriate sub-topics.

Section A focuses on the rationale of conducting the project and what is already known about the topic. As professionalism is a dynamic and socially constructed concept, it was necessary to conduct a thorough literature review to ensure the full understanding of the concept.

In Section B, the focus is on adapting and operationalising this understanding of professionalism in order to develop a valid measure of the professionalism culture. The validity of the instrument was determined by achieving consensus among the primary stakeholders of medical professionalism.

Section C focuses on the outcome of the first field test of the measure which was carried out to determine its reliability and acceptability. This section identifies the strengths and weaknesses, and the changes emanating from the study.

Section D concludes the thesis with a discussion on the overall observations on the process of developing and validating the measure of institutional professionalism culture.

Table of Contents

List of tables

List of figures

List of abbreviations

ABIM	: American Board of Internal Medicine
AMA	: American Medical Association
AMEE	: Association of Medical Educators in Europe
ASME	: Association for the Study of Medical Education
CVI	: Content Validity Index
DBIP	: Dundee Barometer of Institutional Professionalism
DREEM	: Dundee Ready Educational Environment Measure
EFA	: Exploratory Factor Analysis
GMC	: General Medical Council
KMO	: Kaiser-Meyer-Olkin
KS	: Kolmogorov-Smirnov
Mini-CEX	: Mini Clinical Evaluation Exercise
ONS	: Office of National Statistics
OSCE	: Objective Structured Clinical Examination
PASW	: Predictive Analytic Software
PBC	: Perceived Behavioural Control
PCA	: Principal Component Analysis
P-MEX	: Professionalism Mini Evaluation Exercise
PSCOM	: Pennsylvania State University College of Medicine
RCP	: Royal College of Physicians
RCPSC	: Royal College of Physicians and Surgeons, Canada
SP	: Standardized Patient
SPSS	: Statistical Package for the Social Sciences
TACT	: Time, Action, Context, Target
TPB	: Theory of Planned Behaviour
TRA	: Theory of Recent Action
UK	: United Kingdom
UREC	: University Research Ethic Committee
URL	: Uniform Resource Locator
USA	: United States of America
WCC	: White Coat Ceremony

Abstract

Background: Professionalism, in today's context, is a determinant of fitness-to-practise of doctors. Many professional and governing bodies around the world have emphasised the need for educating medical students explicitly about professionalism. In fostering professionalism, the institutional culture plays a concealed but vital role. Although the institutional professionalism culture should be explored and understood there was no suitable measure for use in the context of UK undergraduate medical education. The aim of this project was to develop a valid, reliable and practical measure of institutional professionalism culture.

Methods and results: The project was conducted in two phases. In the first phase, forty six attributes of professionalism were identified in a literature review. These attributes were surveyed among a nationally representative quota sample of 954 members of the UK general public. They identified 44 attributes as important. With a principal component analysis, three facets to professionalism were identified: the relationship of doctors with patients (clinicianship) and co-workers (workmanship), and the behaviour of doctors in society (citizenship). By analysing the survey responses of 368 UK medical professionals using the Content Validity Index, 28 attributes were identified to represent each facet (clinicianship 10, workmanship 11, and citizenship 7).

In the second phase, the 28 attributes were included in an online measure (Dundee Barometer of Institutional Professionalism) with a rating scale based on the Theory of Planned Behaviour (TPB), and field-tested among the faculty and students of the Dundee

Medical School. Based on the TPB, questions on personal attitude, institutional expectation and achievability in relation to each attribute were included. The field-test received 212 responses. The measure demonstrated high internal consistencies at both measure and facet levels. It appeared that the professionalism culture in Dundee Medical School was patient-centred, teamwork-oriented and society oriented. A principal component analysis helped reduce the number of items to 15 with five attributes representing each facet. A generalisability study predicted a highly acceptable reliability with the 15 items. The reaction of respondents towards the measure was positive.

Conclusions: The Dundee Barometer of Institutional Professionalism (DBIP) is the first quantitative measure of the culture of professionalism in UK medical schools. It was developed with the consensus of both professionals and the general public, and used a theory-based rating scale (hence high validity). It is shown to be reliable with 15 items. The DBIP is a practical measure as it is easy to administer and is acceptable to respondents. The construct validity of the DBIP and its ability to distinguish differences in professionalism culture are areas of future research.

Section A: Rationale and background

In this section, the reasons for conducting the research (Chapter 1) and what is already known about the topic (Chapter 2) are discussed.

Chapter 1 – Introduction and rationale

Until recently, professionalism was an area of medical education which has not been 'taught' to students but rather 'caught' by them (Wear & Castellani 2000). Although the importance of incorporating professionalism formally into medical education has been discussed by many for years (Cruess & Cruess 1997; Swick *et al.* 1999), only a few medical schools have implemented it. Certain adverse patient outcomes (e.g. the Shipman case in the UK), which shocked the general public and generated a large public outcry, made the authorities realise the necessity of reforming medical education including the development of professionalism among doctors (Irvine 2005). As a result, professionalism is now recognised as an explicit component in medical education, not only in the UK (GMC 2006) but also around the world (ABIM 2000; RCPSC 2005; Zaini *et al.* 2011). It is argued that 'professionalism needs to be woven into the fabric of the entire curriculum' (Goldie 2008, p.513). In order to achieve this, several outcome models e.g. the three circle model of educational outcomes (Harden 1999), have been developed, and several medical schools in the UK (e.g. University of Dundee), have adopted them (Dundee Medical School 2008). Therefore, contributing to professionalism education has now become crucial in medical education.

Assessment of professionalism at the institutional level, i.e. professionalism culture, which is the main focus of this project, has been emphasised repeatedly as a part of professionalism education for several reasons.

- Educational theories, especially Bandura's Social Cognitive Theory (SCT) (1989), highlight the impact of the institutional environment on learning. According to SCT, people learn behaviours by observing others which is referred to as modelling (Bandura 1989). Therefore, social context plays an important role in gathering new information and developing new behaviours (Bandura 1989). The members of an educational institution of medicine form a micro-society, which is shown to be influential in developing professionalism among undergraduate students (Glicken & Merenstein 2007; Karnieli-Miller *et al.* 2010; Quaintance *et al.* 2010; Wilkes & Raven 2002). On the other hand, the behaviours, which are especially influenced by social learning include certain habits (e.g. reading), qualities (e.g. bravery in adverse situations), reactions (e.g. aggression), and moral judgements (e.g. making decisions on 'good' or 'bad') (Ormrod 1999).

- It is also evident that a significant proportion of professionalism is learnt by students through the 'hidden curriculum' (Hafferty 1998; Swick *et al.* 1999). A major component of the hidden curriculum, i.e. what students learn implicitly rather than explicitly, is represented by the culture of the educational institution and changes made to the institutional culture cause observable effects on students' achievement of expected educational outcomes (Wren 1999).

- One of the theories of moral development describes it taking place in three stages; pre-conventional (morality is governed by rewards or punishments), conventional (morality is governed by expectations), and post-conventional (morality is governed by autonomous principles) (Kohlberg 1973). In medical education, the evidence suggests that the transformation of medical students from the conventional stage to the post-conventional stage has been a consistent problem with some studies demonstrating that there can be an 'erosion' of morality from the third year of undergraduate training with exposure to the clinical environment (Branch 2000; Hojat *et al.* 2009b). An adverse professionalism culture in clinical environments has been blamed for this negative trend (Branch 2000; Hojat *et al.* 2009b). Together with the previous point, professionalism culture appears to affect the formation of professionalism identity among students.

- Brazeau *et al* (2010), who compared empathy, burn out and professional climate, conclude that an adverse professionalism culture is directly related to poor empathy and early burnout among medical students.

- From a sociological perspective, the assessment of professionalism of individuals is ineffective without taking into account the professionalism culture, as professionalism is 'a distributed rather than an individual feature of team function' (Martimianakis *et al.* 2009, p.835)

- In the UK, the importance and necessity of exploring the institutional culture of professionalism as a part of fostering correct professional attitudes and behaviours have been repeatedly emphasised (Irvine 2001; Stephenson *et al.* 2006).

Accordingly, measuring and understanding the professionalism culture at institutional level is important as: the professionalism culture itself is a source for professional development; it is highly influential in the formation professional identity as a doctor; and it affects the attrition of students from medical educational programmes.

Professionalism, however, appears to vary between different stages of medical training, e.g. undergraduate, postgraduate and continuing medical education (Hilton & Slotnick 2005). Therefore, the interventions need to reflect the stage on which they are focused. The professionalism culture of medical schools lays the foundation for fostering professionalism as professionalism lapses gone undetected at this stage may lead fitness-to-practice issues among practising clinicians (Papadakis *et al.* 2001). Therefore, it is timely and educationally rational to explore and measure the professionalism culture of medical undergraduate programmes in the UK, which is the primary aim of this project.

Chapter 2 – Literature Review

As the primary aim of this doctoral project was to explore and measure the professionalism culture of medical undergraduate programmes in the UK, a literature review was conducted to achieve the following objectives:

- to identify the emergence of the concept of 'professionalism' in the history of medicine and medical education as professionalism appears to be a dynamic concept;
- to define professionalism in the current context as professionalism is context dependent;
- to explore professionalism as a part of formal undergraduate curricula as the events and activities of the formal curriculum need to be understood clearly before exploring the informal and hidden aspects; and
- to explore institutional professionalism culture as a part of informal and hidden aspects of professionalism curriculum as professionalism culture appears to be influenced mainly by the informal and hidden curricula.

It was necessary to examine the literature in relation to all the above aspects as professionalism is a dynamic, complex and controversial area in medical education. A comprehensive understanding of the concept was required before the development stage. Therefore, it was also considered as a key part of the data collection process.

In this chapter, the literature is discussed under each of these objectives. The insight generated was used to formulate the research questions and the objectives of this project.

The literature search for this thesis was carried out using the databases: PubMed, Web of Science, EBSCO, and Google Scholar. In addition, the leading journals in the field of medical education namely, *Medical Education, Medical Teacher and Academic Medicine,* were purposefully searched for relevant articles. The key words used are mentioned under each section. Where appropriate, the relevant sections of books written on the topics were referred to.

2.1. Emergence of the concept 'professionalism' in medicine

This section discusses how medicine evolved as a profession and its relationship with the development of professionalism as an explicit component.

In the literature search, key words used were primarily focused on but not confined to; medical, profession, professionalism, definition, conceptualisation.

2.1.1. Medicine as a profession

*'*Several of what we now call *professions* first became organised as guilds*'* and they controlled their own standards of performance (Sox 2007, p.1533). Evidence of forming guilds by doctors dates back to the late medieval and the early renaissance periods, and doctors may have begun to organise as a profession about seven to eight centuries ago (Sox 2007). The social recognition of doctors as professionals, however, was further delayed as they had very little to offer to the general public due to weak knowledge on disease mechanisms and lack of formal medical education (Sox 2007). In the mid-19th century, physicians became the exclusive purveyors of the advancing scientific medical knowledge, 'which made them increasingly indispensable' (Sox 2007, p.1543). This may

have initiated the formation of the professional identity of doctors. Some sociologists, however, argue that medicine transformed from mere occupation to profession only in the 1960s after the introduction of restrictive licensing laws for people who could practise medicine, especially in north America (Hafferty & Castlellani 2009; Martimianakis *et al.* 2009) Nevertheless, it is evident that the significant factors in the development of medicine as a profession were the possession of specialist knowledge, the recognition of its uniqueness by the public, and the emergence of regulatory frameworks.

'Profession is a term used to describe many occupations (e.g. law, the clergy, accounting, architecture and engineering) and is therefore generic' (Cruess *et al.* 2004 ,p.75). The literal definition of *profession* as a paid occupation, which requires advanced education and training (Hornby 1995), however, may not satisfactorily convey its broad meaning. At very least, the word *profess* in the Hippocratic Oath referred to a public commitment to a set of values (Cruess *et al.* 2000b). Citing Brandeis (1912), Arnold and Stern (2006) propose that 'profession' can be defined as providing service to society after necessary preliminary training, which is pursued largely for others and not merely for oneself, and in which the amount of financial return is not the accepted measure of success. While echoing Brandeis's thoughts Cruess *et al* (2004) broaden the definition of *profession* with the elements of self-regulation and the governance by ethical codes, based on historical and sociological observations. *A profession,* therefore, is a self-regulated occupation which requires specialist knowledge and skills, and with a primary responsibility for serving society ethically. This implies that the functions of a profession could be viewed from two perspectives: self-preservation, i.e. taking collective actions such as setting standards of

practice; and an individual's actions and behaviours, which are required to enhance societal well-being (Sox 2007). Although there are commonalities among various professions, the unique service provided to the society by each professional group (e.g. physicians, lawyers, architects) determines the nature of professional attributes and behaviours (Cruess *et al.* 2004). Therefore, in the context of this thesis, the definition of *professionalism* requires to be narrowed down to medicine or, at least, to healthcare professions generally.

2.1.2. Professionalism in medicine

Professionalism is discussed only rarely in most of the medical education literature up to the 1970s (Arnold 2002). Arnold and Stern (2006) argue that 'although humanism and definitions of "professionals" have long histories, the word "professionalism" as currently used is a recent phenomenon in medicine' (p.18). However, they do not negate the presence of an interest in certain characteristics expected from physicians and students, which were regarded as a residual category referring to qualities that were not cognitive, in the recent past (Arnold & Stern 2006). Sociologists directly refer this residual category as professionalism, maybe because their point of view and conceptualisation of medical professionalism is different to doctors' conceptualisation of professionalism (Hafferty & Castlellani 2009). From the sociological perspective, Martimianakis *et al* (2009), argue with evidence that doctors, in particular, have been talking about professionalism for about 100 years. Therefore, examination of the concept of professionalism in the past and appreciation of changes over time (if there are any) is relevant and worthwhile in measuring professionalism (Martimianakis *et al.* 2009). It appears that, historically,

professionalism has been conceptualised and operationalised in several forms, which are discussed below, quoting the major developments as examples.

2.1.2.1. Professionalism in the early stages of medicine

The original Hippocratic oath for medical practitioners, 'perhaps the most widely known of Greek medical texts' which is estimated to be written in the 4th century B.C. by an unknown author (National Library of Medicine 2009), may be the first of its kind. The main elements were: teaching medicine (which is referred to as an art rather than a science) to the next generation; making decisions for the betterment of patients; doing no harm or injustice; conforming to moral principles, social norms and laws; accepting limitations and respecting collegiality; not exploiting one's position; and keeping the confidentiality of information provided by patients (National Library of Medicine 2009). This oath largely focused on the noble nature of the profession and reflects the need for developing medicine as a profession rather than as a vocation (Sox 2007). A vocation requires a qualified person and a profession requires a qualified person with the ability to self-regulate (Cruess *et al.* 2004).

2.1.2.2. Professionalism in the 19th century

The notion that professionalism is responsive to social and political changes is obvious in the code of conduct published by the American Medical Association (AMA) in 1847, an era in which the industrial revolution was gaining momentum (Sox 2007). Unlike in the Hippocratic Oath, teaching the next generation of doctors was not emphasised in the AMA code. As doctors were the sole stakeholders of growing medical knowledge and medico-

technical capabilities in this era, they may have prioritised commercial and research interests over patients' interests. This is reflected in the code by the assimilation of elements such as: not advertising; not subjecting patients to unnecessary medicine or instrumentation; offering care for family members of physicians without charge; not meddling in the care of other physicians' patients; exposing quackery; and avoiding excessive visits to the patient (Sox 2007).

2.1.2.3. Professionalism in the 20th century

The original Hippocratic Oath, however, was adopted and transformed to modern-day medicine by Louis Lasagna in 1964. It has been used in many medical and dental schools around the world (Nova Online 2009). The main elements in this version include: teaching medicine (which is referred to as both a science and an art) to the next generation; making decisions for the betterment of patients; doing no harm or injustice to patients; prescribing treatments rationally; understanding social and economic impact on patients in the treatment process; observing patients' rights; accepting one's limitations; respecting collegiality; keeping the confidentiality of information provided by patients; and promoting disease prevention (Nova Online 2009). In the new version, there was no emphasis on conforming to moral principles or social and legal etiquette. Unlike the 1847 AMA code, the modern version of the Hippocratic Oath does not focus largely on preventing commercialisation of the profession. Instead, the scientific basis of medicine, as opposed to viewing medicine as an art in the older version, and its use in prescribing treatment taking into account patients' social and economic concerns have been

emphasised. The focus on protecting patients' rights is a major development in the modern Hippocratic Oath.

2.1.2.4. Professionalism in the 21st century

The AMA code was modified in 2001 in order to meet the socio-political changes of healthcare delivery (Sox 2007). The AMA 2001 version expects physicians to be dedicated in providing competent patient care with compassion and respect for human dignity and rights (Sox, 2007 citing AMA House of Delegates, 2001). In addition, it emphasises the necessity of doctors: being honest; reporting colleagues who are incompetent, fraudulent or deceptive to relevant authorities; conforming to laws; respecting the rights of patents, colleagues and other healthcare professionals; protective of patients' confidentiality within legal limits; participating in the advancement of medical science and continuing medical education, and enhancement of health education among the general public; working with colleagues and other healthcare professionals for the best interests of patients; being non-discriminatory, responsible for health promotion, altruistic, and supportive of improving access to medical care for all people. The primary focus of professionalism in this code, therefore, is adopting a patient-centred approach in the practice of medicine, while being a colleague, a member of the health care team, and a citizen. The Physician Charter produced by the American Board of Internal Medicine in 2002 widened the horizon of medical professionalism in the USA (Sox 2002). It has furthered the patient-centred approach in the practice of medicine mentioned in the AMA code 2001, and identified professionalism as the basis of medicine's contract with society

(Sox 2002). The charter is based on three fundamental principles: primacy of patient welfare, patient autonomy, and social justice (Sox 2002).

In 1993, the General Medical Council (GMC), UK, introduced and defined professional attitudes expected from graduating doctors in their first edition of *Tomorrow's Doctors* (GMC 1993). However, professionalism in medicine has been specifically emphasised in its subsequent editions (GMC 2003, 2009), and with the introduction of guidelines for practising doctors in *Good Medical Practice* (GMC 2006) and for medical students in *Professional Behaviour and Fitness to Practice* (Medical School Council 2007).

The GMC guidelines on good medical practice are all about setting out 'the principles and values on which good medical practice is founded'; these principles together describe medical professionalism in action at least as perceived in the UK context (GMC 2006). In both *Good Medical Practice* and *Professional Behaviour and Fitness to Practice,* the principles of professionalism described are: good clinical care; maintaining good medical practice; teaching and training, appraising and assessing; relationship with patients; working with colleagues; probity; and health. Sub-areas have been specified under some of these principles (Table 1). However, these principles have been criticised by some, arguing that they are a misinterpretation of professionalism; these principles are a set of duties, breaching which is punishable, whereas professionalism is a set of moral obligations (Horton 2005).

Table 1- Principles of professionalism as described in *Good Medical Practice* (GMC 2006)

Domain	Attributes
Good clinical care	Providing good clinical care
	Supporting self-care
	Avoid treating those close to you
	Raising concerns about patient safety
	Decisions about access to medical care
	Treatment in emergencies
Maintaining good medical practice	Keeping up to date
	Maintaining and improving your performance
Teaching and training, appraising and assessing	
Relationships with patients	The doctor-patient partnership
	Good communication
	Children and young people
	Relatives, carers and partners
	Being open and honest with patients if things go wrong
	Maintaining trust in the profession
	Consent
	Confidentiality
	Ending your professional relationship with a patient
Working with colleagues	Working in teams
	Conduct and performance of colleagues
	Respect for colleagues
	Arranging cover
	Taking up and ending appointments
	Sharing information with colleagues
	Delegation and referral
Probity	Being honest and trustworthy
	Providing and publishing information about your services
	Writing reports and CVs, giving evidence and signing documents
	Research
	Financial and commercial dealings
	Conflicts of interest
Health	

In summary, the review of the literature in this section revealed that professionalism is the characterisation of medicine as a profession that has evolved with socio-economic changes and societal expectations. For example, in the early ages of medicine, the emphasis of professionalism was on the transformation of the vocation of medicine into a noble profession. This was strengthened further during the industrial revolution by protecting the profession from commercialisation. In the 20th century, professionalism was about establishing patient-centred medical practice by emphasising evidence-based practice and patient safety. By the 21st century, this has been transformed virtually into a contract between doctors and society to ensure rights of patients and to encourage their trust towards the practice of medicine. Therefore, this dynamicity of the concepts should be considered in defining professionalism for the purpose of this study and any intervention should reflect the current focus of professionalism.

2.2. Definition of professionalism

As discussed above, professionalism is a dynamic concept. Therefore, it is necessary to work out a definition or a conceptual framework which is appropriate for the time and context.

In the literature search key words used were primarily focused on but not confined to; medical, profession, professionalism, definition, conceptualisation.

2.2.1. Perspectives on professionalism

Medical professionalism in the 21st century has been defined or conceptualised in the perspectives of medicine and sociology, which reflect the differences in points of view

between the two sciences (Hafferty 2006). The common feature of medicine's own definitions of medical professionalism is that professionalism is considered largely as a competency or an outcome, which is necessary to establish a sound doctor-patient relationship (Hafferty 2006; Martimianakis *et al.* 2009). This concept is put forward by many professional and regulatory bodies and most of the medical literature represents this perspective (Hafferty 2006). For example, the '*Good Medical Practice*' of the GMC (GMC 2006) (Table 1) and the professionalism framework of the American Board of Internal Medicine (ABIM) (ABIM 2000) (Table 2) have identified distinct aspects of professionalism. As evident in the literature, a similar approach has been taken by many medical educationalists (e.g. Lambe & Bristow 2010; Rogers & Ballantyne 2010; Sehiralti *et al.* 2010).

Table 2 - Domains of professionalism as defined by American Board of Internal Medicine (ABIM)

Altruism
Accountability
Excellence
Duty
Honour and integrity
Respect for others

From the sociological perspective, medical professionalism is socially constructed and cannot simply be reduced to traits and behaviours (Hafferty & Castlellani 2009; Martimianakis *et al.* 2009). Therefore, sociologically, professionalism is the result of professionalisation, a process in which the professional identity of doctors is determined not only by his/ her qualities but also by multiple factors such as his/her response to needs and demands, social norms, behaviours of people around them and organisational

structure (Martimianakis *et al.* 2009). However, from the point of view of many medical educationalists, for professionalism to be fostered, encouraged and assessed, it should be operationalised in some form of observable acts (Hoff 2000; van Mook *et al.* 2009c). Therefore, in any assessment or measurement on professionalism, professionalism should be defined as traits and behaviours (complying with medical perspective) but should also be understood and interpreted with due consideration to the underlying social context (complying with sociological perspective).

2.2.2. Constituents of professionalism

The constituents of professionalism are still a matter of debate that has resulted in multiple definitions and conceptualisations (Arnold 2002; Wilkinson *et al.* 2009). Hornby (1995), for example, defines professionalism as the knowledge, skills or qualities required or expected of members of a profession. However, after a review of the literature, Rogers and Ballantyne (2010) conclude that both clinical competence, i.e. knowledge and skills, and ethico-legal competencies should be outside the scope of professionalism *per se*. Both the GMC and ABIM frameworks reflect this conclusion of Rogers and Ballantyne (2010), as both ethical and clinical competencies have been considered as the basis for, but not as a component of professionalism. Cohen (2007), on the other hand, claims that humanism is a related but a separate construct from professionalism. However, these differences are not universal as several definitions of professionalism include moral, ethical and legal principles, and humanistic qualities.

The attributes of professionalism identified from the literature range from interpersonal and personal skills (e.g. teamwork, communication, reflective practice) to values (e.g. honesty, integrity, accountability), and attitudes (e.g. collegiality, respecting patients' autonomy). These are discussed more comprehensively in Chapter 3 (p.110).

2.2.3. Working definition

Epstein and Hundert (2002) define professionalism as 'the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served' (p.226). This definition broadly encompasses the constituents of professionalism proposed in many other definitions (Cruess *et al.* 2004; Hilton & Slotnick 2005; Jha *et al.* 2006; Swick 2000; van-de-Camp *et al.* 2004; Wagner *et al.* 2007), and allows for the evolution of professionalism, i.e. its variability with societal expectations from time to time (Cruess *et al.* 2004), and cultural background (Cruess *et al.* 2010; Hafferty 2006). However, it does not take into account the perspective of the sociologists discussed above. Therefore, the definition by Epstein *et al* (2002) is modified as follows to be used as the working definition of this project.

'Professionalism is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice in a given context for the benefit of the individual and community being served as normatively defined by all stakeholders.'

The contextualisation and the appreciation of the role of stakeholders in setting norms will satisfy the sociologists' perspective of professionalism as it reflect the socially constructed nature of the phenomenon.

2.3. Exploring professionalism as part of the formal undergraduate curriculum

Professionalism is equally important across the continuum of medical education, i.e. undergraduate, postgraduate and continuing medical education. However, at least certain components of professionalism appear to be stage-specific (Hilton & Slotnick 2005). Therefore, complying with the focus of this project, the stage of undergraduate training is mainly considered in the discussion below.

With the enhanced understanding of what is meant by 'professionalism' in medicine in the previous section, this section now explores how professionalism is taught and learned, i.e. fostered, and assessed, specifically at the undergraduate level on which this project is focused.

2.3.1. How is professionalism fostered?

The following questions were answered with literature evidence to understand how professionalism can be fostered.

a. Can professionalism be fostered? (Professionalism as an educational construct)
b. How are medical curricula designed for professionalism education? (Curriculum designs for professionalism education)
c. What methods have been used to foster professionalism? (Methods used in fostering professionalism)

d. How should professionalism be taught? (Single versus multi-prong approach)

In the literature search key words used were; medical, professionalism, professional behaviour, professional attitudes, undergraduate, teaching, learning, fostering, curriculum.

2.3.1.1. Professionalism as an educational construct

The general agreement is that the ultimate goal of professionalism education is supporting the formation of an identity of professionalism in students, i.e. the development of the distinct entity of a 'good doctor' (Hafferty 2006; Monrouxe 2010; van Mook *et al.* 2010). Professionalism primarily encompasses one's virtues and morality (Bryan & Babelay 2009; Remen & Rabow 2005). It is more than just another educational competence and may involve the modification of existing moral outlooks of students (Huddle 2005; Monrouxe 2010). Aristotle believed that 'morality cannot be taught but needs to be practised' (Grodzinsky 1999). Although there is still some evidence to support the view that such attributes cannot be taught as they are inborn (Finn *et al.* 2009; Knights & Kennedy 2006), the acceptance of this argument is gradually losing ground, as the evidence lacks specificity on the attributes that cannot be taught and more evidence is emerging to suggest the contrary (Cruess & Cruess 1997; Hodges *et al.* 2011). Many regulatory and governing bodies around the world have emphasised the attainment of professionalism more as an exit outcome than as a selection criterion (e.g. ABIM 2000; Medical School Council 2007; RCPSC 2005) which implies that professionalism can be fostered during undergraduate education. A growing number of medical educationalists

take the stance that professionalism is an acquired state rather than a trait (Cruess & Cruess 1997; Duff 2004; Hilton & Slotnick 2005; Wear & Castellani 2000). This view has been backed by psychological research. Reviewing the work of many psychologists and their on-going research, Bryan and Babelay (2009) conclude that virtues of practical wisdom (e.g. critical thinking, ingenuity, open-mindedness), justice (e.g. conscience, duty, ethics, honesty, respect, teamwork), temperance (e.g. self-control, prudence, discretion), fortitude (e.g. altruism, integrity), faith (e.g. spirituality, trust), hope (e.g. future-mindedness, optimism), and love (e.g. beneficence, compassion, empathy, kindness, politeness) can be fostered, at least to some extent.

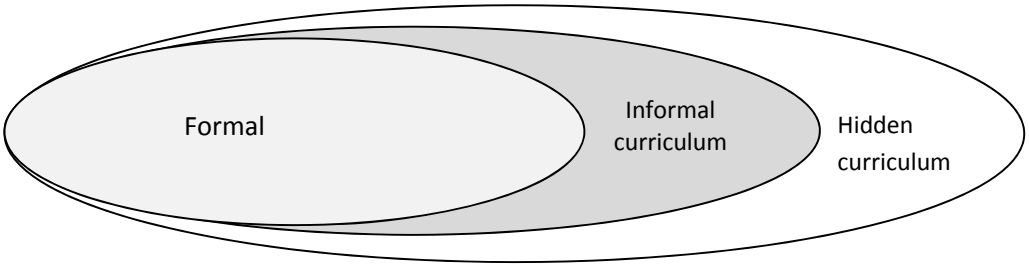
Both research (e.g. Branch 2010; Bryan & Babelay 2009) and the experiences of medical educationalists (e.g. Goldie 2008; Hilton & Slotnick 2005; Kumagai & Lypson 2009; Mueller 2009) suggest that 'teaching' should cross traditional boundaries to foster necessary professional attributes among students, which makes the task challenging (Jones *et al.* 2004). Despite the challenges, professionalism education should be emphasised as lapses in professionalism among junior doctors appear to be a significant issue. For example, in 2002, the majority of US clerkship directors viewed that the communication skills and professionalism of medical graduates were unacceptable (Windish *et al.* 2004).

2.3.1.2. Curriculum designs for professionalism education

In professionalism education, not only 'what is taught' (the intended components), but also 'what is caught' (the subtle messages emerging from peoples' behaviours and institutional culture) matter (Cruess & Cruess 2006; Goldie 2008; Monrouxe 2010).

Hafferty (1998), therefore, has proposed that, in a given learning environment, three curricular aspects, formal, informal and hidden, operate to form the professional identity of students. The formal curriculum is what is planned, endorsed and offered by the educational programme; the informal curriculum is what students learn in an ad hoc manner by interacting with teachers as a part of the teaching / learning process; and the hidden curriculum is the culture and morality portrayed by the members of the institution as a community (Hafferty 1998). In many instances, both informal and hidden aspects of the curriculum contradict the formal curriculum (Hafferty 1998). In the perspective of 'learning', the relationship between the three aspects of curricula in the development of professionalism is illustrated in Figure 1.

Figure 1 – Aspects of 'professionalism' curriculum



The impact of the informal and hidden curricula on fostering professionalism appears to be greater than the formal curriculum (Cruess & Cruess 2006; Ginsburg *et al.* 2003; Hafferty 1998). The professionalism culture of an institution is reflected considerably in its informal and hidden curricula (Hafferty 1998). Therefore, the institutional culture appears to be a primary determinant in fostering professionalism among students (Hafferty 1998). The importance of the formal curriculum in professionalism education, however, cannot

be underestimated (Cruess & Cruess 2006; Kalet *et al.* 2002). In fact, the formal approaches should strive to encroach into the informal and hidden curricula with constant effort at understanding the latter (Ginsburg *et al.* 2003; Kalet *et al.* 2002). In a study conducted in England, Wales and Australia, Monrouxe *et al* (2011), conclude that the 'opportunities [for students] to engage in active sense-making activities within the formal professional curriculum can encourage an embodied and sophisticated understanding of professionalism' (p. 585). Therefore, the main focus of this section is on understanding the formal approaches of fostering professionalism. As the measurement of the professionalism culture in an institution is the focus of this doctoral project, it is discussed in length elsewhere in this literature review (p.61).

a. Principles guiding the formal professionalism curriculum

Hilton and Slotnick (2005) propose that approaches used to foster professionalism at various stages of undergraduate training should reflect the psychological and moral development of students, and should facilitate their journey from naivety to practical wisdom (*phronesis*) and discourage cynicism over idealism. This proposal has been made relevant by several studies. Madill and Latchford (2005), in a study conducted with first-year students at Leeds Medical School concluded that, even within the first year, students acquire certain professional attributes of practical wisdom, e.g. dedication, competence, responsibility and coping strategies, while losing personal attributes such as self-confidence and enthusiasm for learning, which may lead to cynicism. Boenink *et al* (2005) compared the performance of Dutch pre-clinical and clinical students in professionalism case vignettes and found that the understanding of professionalism among the clinical

group was better than the non-clinical group probably due to increased socialisation of the former in clinical environments.

In achieving the complex educational goal of fostering professionalism,

- Professionalism education should be thoroughly integrated both vertically and horizontally throughout the curriculum (Stirrat *et al.* 2010). For example, students of Liverpool Medical School identified that their traditional, discipline-based curriculum was especially weak in fostering professionalism (Watmough *et al.* 2009).
- Outcome-based education has been suggested as an effective concept for curriculum design as it equally emphasises both the technical and professionalism aspects in an integrated manner (Goldie 2008). However, the effectiveness of this approach at the implementation level is questionable. For example, 'professionalism' has still been given lesser emphasis than technical competencies in US medical schools (Mooney *et al.* 2010).
- Cruess and Cruess (2006) propose that the theory of situated learning is appropriate as the basis for professionalism education. The situated learning theory 'suggests that learning should be embedded in authentic activities, which help to transform knowledge from the abstract and theoretical to the usable and useful' (Cruess & Cruess 2006 p.205).
- Unlike the traditional model of teaching ethics (teaching ethics mainly by university staff, who may not always be practicing clinicians), professionalism education should be a shared obligation of both university and clinical teachers, and the process of

professional socialisation should be encouraged by multi-professional groups (Buyx *et al.* 2008; Stirrat *et al.* 2010).

b. Design and arrangement of formal professionalism curriculum

There are reports in the literature of dedicated courses and / or existing courses with or without modifications used as single or serial interventions to foster professionalism.

b.1. Examples of the use of dedicated courses for professionalism education

Dedicated courses on professionalism have either been offered to students as components of the core curriculum or as optional / elective components.

- Professionalism as a core curricular component

In the USA, Makaul and Altman (2002) used students' encounters with standardised patients (SPs) to teach 'informed consent' to first-year students. At the end of the encounters, the SPs provided students with feedback. As the encounters were video-recorded students were also given the opportunity to self-reflect. The students who participated in the intervention evaluated it very positively. This educational intervention, therefore, encourages experiential learning and promotes reflection. SPs were used to encourage empathy during consultations by Wagner *et al* (2002) in the USA and Towle and Hoffman (2002) in Canada, but in a slightly different way. In their approach, the relevant principles and concepts were introduced to students in small groups before exposing them to SPs. In the introductory component, Wagner *et al* (2002) used the analysis of a set of purpose-made videos while Towle and Hoffman (2002) used pre-

reading supplemented with video demonstrations. After the subsequent encounters with SPs, students received feedback from both SPs and clinicians. The respective student groups liked the intervention and they perceived that it helped them understand their professional role as well as develop their confidence in carrying out doctor-patient consultations. The University of East Anglia, UK, adopted a longitudinal consultation skills education course that incorporated several professional attributes such as patient-centredness and identifying own limitations (Papageorgiou *et al.* 2011). They used lectures, problem-based learning sessions and SPs for teaching. The students perceived the whole course, the experiential learning components in particular, very positively. In both these interventions cognitive inputs as well as experiential learning were used to provide the necessary knowledge and encourage reflection. The four-week communication course introduced by Ang (2002) for fourth year Chicago medical students specifically focused on the management of conflict of interest. She provided a structured framework, small group discussions, role plays and reflective discussion on role plays, which were received well by participating students; the course helped them develop confidence and positive attitudes towards conflict management. In this intervention, Ang (2002) used cognitive inputs and experiential learning. A spirituality training course was introduced to third- year students at the University of Missouri-Kansas City for helping students understand patients' spiritual beliefs and for using spirituality to improve doctor-patient relationships (Graves *et al.* 2002). In addition to a lecture (cognitive inputs) to introduce the term 'spirituality', the course organisers used small group discussion (experiential learning) and shadowing sessions with university chaplains (in-practice

learning). It is interesting to note that students liked the component of in-practice learning the most and this component reduced the doubts and resistance among students regarding the relevance of the topic to their professional development.

- Professionalism as an optional / elective curricular component

In the Massachusetts Medical School, a week-long elective was conducted with both didactic, group and in-practice activities in classroom environments to develop teaching skills of fourth year students (Pasquale & Pugnaire 2002). In the University of Arkansas, the clinical skills centre was used to develop students as teachers (Moseley *et al.* 2002). The students have reported that these courses provided them with necessary knowledge and skills. The 'healers' art' is an elective course with five three-hour modules, offered by 33 medical schools in North America with the objective of inculcating moral obligations and values (Remen & Rabow 2005). The modules use staff presentations, guided reflections on personal experience and small group discussions (Remen & Rabow 2005). On average, the course was subscribed to by 44% of the student population. They rated the course positively in terms of its content and delivery (Remen & Rabow 2005). Literature courses have been used primarily to enhance humanistic aspects of medicine, e.g. empathy, altruism, compassion, and caring toward patients (Shapiro & Rucker 2003). In a case-control study, Shapiro *et al* (2004) demonstrated that a short (eight one-hour sessions) literature course (on-site reading of poetry, skits and short stories) focused on the doctor-patient relationship, physical examination, listening to patients, pain, sexuality, cross-cultural issues, lifestyle modification/noncompliance and geriatrics, offered as an

elective to the first year students of a Californian medical school, significantly improved their empathy. Although there are benefits, literature and related courses, e.g. arts and music, should not be made compulsory for students. According to Gordon (2003) imposition of such courses on unwilling students, who would learn professionalism more effectively by alternative means, may make them become demoralised.

b.2. Examples for the use of existing courses for professionalism education

The importance of basic science courses in fostering professional attributes among young medical students has been repeatedly emphasised (Pawlina 2006; Swartz 2006; Swick 2006). The medical students of Ulm University, Germany, perceived that their anatomy dissection course offered them opportunities to learn team-work, coping strategies and time-management, though the course had little to offer in fostering empathy (Bockers *et al.* 2010). In the University of California, students' understanding of professionalism was improved by introducing an additional creative component to a dissection course, in which students were expected to reflect upon anatomy dissections by means of literature, music or media (Shapiro *et al.* 2009). At the Philadelphia College of Osteopathic Medicine, video tutorials related to doctor-patient relationship were introduced in parallel with the basic sciences curriculum which helped incorporate professional values (DiLullo *et al.* 2009). Although the course was non-compulsory most students made use of the tutorials. In the Mayo Clinic Medical School, video interviews with family members who donated cadavers were offered to pre-clinical students with the view of enhancing the understanding of professionalism, including the attributes of confidentiality, respectful

behaviour and humanism in medicine (Kostas *et al.* 2007). According to the feedback of students the intervention met its goal.

After an international health elective programme, Canadian students were encouraged to reflect upon cultural and social differences between well-resourced and poorly-resourced countries in the delivery of healthcare (Elit *et al.* 2011). This intervention helped students improve their cultural understanding and the authors suggest making future students aware of these differences before they take up international electives. The clinical programme for Michigan medical students was changed to provide protected time for professionalism education; students reflect and discuss their clinical experiences in small groups with tutors during these times (Andre *et al.* 2003). Creating fora to share student professionalism experiences in clinical environments is an effective mode of fostering professionalism (Monrouxe *et al.* 2011).

2.3.1.3. Methods of fostering professionalism

Different teaching and learning strategies have been used to impart knowledge, to provide opportunities for experiential learning, and to encourage reflection-on-practice in fostering professionalism among medical students. These strategies may be single interventions or multi-prong approaches. In addition, staff development activities have been geared towards professionalism education to reinforce these teaching and learning strategies.

a. Imparting knowledge related to professionalism

Documents published by regulatory and governing bodies, lectures or plenaries, e-learning interventions and dedicated days / events were the commonly used approaches for imparting the relevant knowledge base related to professionalism.

- *Documents on professionalism*

Guidelines published from time-to-time by professional and governing bodies, e.g. the General Medical Council (UK), are the primary resource for cognitive inputs for medical students (Medical School Council 2007). These frameworks guide students, especially in early years, on professional behaviours (Robins *et al.* 2002). However, they need to be emphasised repeatedly not only to undergraduates (Garner & O'Sullivan 2010) but also to postgraduates and practitioners (Chandratilake *et al.* 2011), as their perceptions may become different to what is expected. As observed in the reflective portfolios of the final-year medical students at the university of Cambridge, the current conceptualisation of professionalism may not have been conveyed effectively to students via the GMC documents; students still believed in some of the elements of 'old' professionalism rather than 'new' professionalism (Borgstrom *et al.* 2010) (Table 3).

Table 3 - The key elements of old and new professionalism (after Borgstrom *et al*., 2010)

Old Professionalism	New Professionalism
Detachment	Empathy
Paternalism	Emotional Engagement
Restricted communication with patients	Open Communication
	Patient-centeredness
Medical beneficence as the most prominent ethical principle	Patient autonomy as the most prominent ethical principle

Similarly, in a national study in the USA, students did not identify professional aspects such as reference to competence, making advances in scientific knowledge, avoiding harm to patients, avoiding sexual misconduct, protecting confidentiality, involvement in disease prevention, or opposing crimes against humanity, which are prominent features of the physicians' oaths in America, as elements of professionalism (Rabow *et al.* 2009). However, they placed importance on aspects like maintaining personal-professional balance, sensing the vulnerability, dealing with personal fears, seeking mutual benefit in patient-physician relationships, supporting colleagues, learning from patients, and achieving personal growth through professional work, which were recognised formally as important aspects of professionalism (Rabow *et al.* 2009). Repeated emphasis on formal requirements in professionalism have been emphasised with the emergence of social media networks, e.g. Facebook (Garner & O'Sullivan 2010) and YouTube (Farnan *et al.* 2008), where many students tend to ignore professional guidelines. A nationwide survey conducted by Chretien *et al* (2009) in the USA revealed that many students post materials which violate patient confidentiality, use profanity and discriminatory language, depict intoxication or are sexually suggestive, on social media sites ignoring existing professionalism guidelines.

The familiarity with guidelines to meet these challenges may be improved by providing opportunities for students to share their opinions with each other and engage in workplace-based activities in light of formal guidelines from their governing bodies (Howe *et al.* 2009).

- Lectures or plenaries

Lectures are another method of imparting the knowledge base related to various aspects of professionalism. They may provide junior students with the professional etiquette for working in the clinical environment (Skochelak *et al.* 2001). However, as observed in one British medical school, they may not be very effective in changing the attitudes and behaviours of students, e.g. in safeguarding confidentiality of patient information, if what they see in practice tends to override what they hear in the classroom (Jethwa *et al.* 2009). In promoting ethical behavior among students, small group discussions have better impact than lectures (Goldie *et al.* 2004).

- Dedicated days / events on professionalism

'White Coat Ceremonies' (WCC) are conducted by many medical schools primarily in the USA for their first-year students with the participation of family members and friends of students (Russell 2002). During the ceremony, students listen to an inspirational talk delivered by an eminent physician, they are physically cloaked with the typical white coat for the first time, and they swear the Hippocratic Oath (Russell 2002). It is meant to create a psychological contract of professionalism and empathy among students and has been recognised as an opportunity to develop a sense of responsibility and trust needed as medical professionals (Russell 2002). However, the usefulness of this ceremony alone to achieve this purpose has been challenged. At the University of Arkansas, a symbolic reminder of integrity, compassion and mastery by means of a badge has been introduced to students and the staff while they work in clinical environments to remind what they

learned from the WCC; students and staff members were happy to wear the badges (Deloney 2003). Similarly, pre-ceremony orientation activities (e.g. preparing oaths) have been attempted with positive student feedback (Fresa-Dillon *et al.* 2004). However, it has been argued that WCC are undesirable as they also inculcate senses of authority and social hierarchy among students which can result in distancing doctors from society (Goldberg 2008; Russell 2002). Cohen *et al* (2002) have attempted to emphasise and retain the intended purposes of the WCC by introducing a small group activity, in which students read around the values expressed during the WCC and develop their own code of conduct. The authors report that the code developed by students emphasised patient autonomy and respect, beneficence, and professionalism. Students not only enjoyed the experience but also learned from it. They perceived that the group-based activity particularly helped them learn professional values. This intervention highlights the benefits of supplementing cognitive inputs such as WCCs or reading materials with experiential learning opportunities (e.g. small group activities) in fostering professionalism.

Although the WCC do not exist in the same form in the UK, oath-taking at graduation ceremonies is not uncommon (Hurwitz & Richardson 1997). The key difference is the WCC are usually held in second or third year, the point at which students enter the clinical phase of the course while the graduation ceremony denotes the end of undergraduate training. 'Oath taking commits doctors to observe an ethical code' (Hurwitz & Richardson 1997, p.1671). However, the contents of traditional oaths, e.g. Hippocratic Oath, may not adequately prepare doctors for new professional obligations and complex moral

predicaments that emerge from scientific and technological advancements (Hurwitz & Richardson 1997).

- e-learning interventions

Roff and Dherwani (2011) introduced an e-inventory which helps students learn expected professional behaviours. In the process of learning, students are expected to engage with the inventory items online, reflect and compare their responses with peers and teachers. Although there is an element of experiential learning, it is largely to make students aware of professional behaviours. McLachlan (2010) introduced a point-based system primarily to provide Durham (UK) medical students with feedback on their professionalism. In this exercise, conscientiousness was used as a measure of professionalism and, routine activities of students (e.g. submission of assignments on time, attendance for compulsory activities) were used to gauge their conscientiousness. Although, the practicability and reliability of this exercise as a measure of professionalism appeared to be high, its acceptability and validity were questioned (McLachlan 2010). Using this exercise as a method of giving feedback to students on certain aspects of professionalism (e.g. acting with responsibility), however, cannot be ignored.

Although important, making students aware of professionalism alone would achieve little in terms of the formation of professionalism among students (Cruess & Cruess 2006). Sharing the personal experience and the programme evaluation findings at the Harvard Medical School, Branch (2010) concluded that effective professionalism education should

also include: longitudinal learning in small groups; creating a supportive group process; prominent inclusion of reflective learning; and experiential learning.

b. Fostering professionalism by experiential learning

Experiential learning denotes the methods, which encourage student learning from their own experiences as learners (Spencer & Jordan 1999). It is characterised by self-directed, inquiry-based and active learning facilitated by teachers (Spencer & Jordan 1999). Experiential learning could take place in several settings, e.g. in classrooms, clinical or community settings.

- Examples of using experiential learning opportunities in the academic institution setting:

Small group discussions and electronic media have been used for experiential learning in classroom settings.

Small group discussions

Small group sessions are particularly important methods of experiential learning. However, they should be stimulating, interactive and honest to achieve desired professionalism related goals (Kumagai & Lypson 2009). Evaluating the Glasgow medical curriculum, Goldie *et al* (2004) concluded that the first-year course, which primarily uses small groups, has a higher impact on students' learning of ethics and legal aspects compared to the later years, which predominantly use lectures. Lypson *et al* (2002) at Michigan and Northwestern Medical Schools in the USA have developed a series of small

group sessions for third-year students facilitated by physicians to encourage reflection on professionalism related issues. In the University of California, an award-winning book has been used as the trigger for similar small group activities (Lie *et al.* 2002). Students have received both courses positively and reported that the courses improved their understanding of professionalism. Short video tapes of real doctor-patient encounters have been used by Ber and Alroy (2002) in Israel as triggers for small group discussions on professionalism for pre-clinical students.

Electronic media

Wikis are collaboratively authored webpages maintained by groups with common needs (Varga-Atkins *et al.* 2010). These web pages create a shared resource through online contributions of their members. Wikis were used experimentally by Varga-Atkins *et al* (2010) in a Problem-Based Learning course at Liverpool with the intention of improving professionalism in peer-to-peer interactions. The students' online interactions were facilitated by academics. The students reported that wikis not only enhanced their knowledge of professionalism but also helped them reflect and challenge their own moral grounds in different professionalism dilemmas. Christner *et al* (2010) from the Michigan University Medical School used a series of emails, pretended to be from patients to teach professionalism to students during their paediatric rotation. Students' responses to these emails were discussed in follow-up group sessions. The performance of individual students was assessed against a valid and reliable marking scheme after each exercise and feedback session. The authors claim that the test scores of students improved significantly over time. The online video tutorials used by the Philadelphia College of Osteopathic

Medicine described above are another example of using e-learning as a method of delivering professionalism courses (DiLullo *et al.* 2009).

- Examples of using experiential learning opportunities in clinical and community settings

In a comparative review, O'Toole *et al* (2005) concluded that, in the USA, cognitive inputs and academic institution-based learning alone have left gaps in professionalism education, which should be addressed by workplace-based learning. Such an approach was supported by the residents and trainees of the University Of New Mexico School Of Medicine, who in hindsight claimed that more practical and workplace-oriented approaches should have been used in teaching professional behaviour during the undergraduate years (Roberts *et al.* 2005). As revealed in a mixed method study conducted by the University of Amsterdam, working in authentic settings with appropriate responsibilities was demonstrated to be highly effective in fostering professionalism among medical students, even as early as in the first year (Scavenius *et al.* 2006). Role models and reflection on practice, however, appear to be the main sources of learning professionalism in clinical and community settings.

Role models

The usual 'workplace' for students is either the academic or clinical environments. Role models are the primary source of learning in both these environments (Baernstein *et al.* 2009; Jones *et al.* 2004). Students challenge or accept the moral grounds behind professionalism-related issues by observing role models (Baernstein *et al.* 2009; Cruess *et al.* 2008). However, the community may be a more appropriate 'workplace' than the

medical school or the hospital in enhancing cultural competence, for example, and the source of learning becomes the members of the community (Kumagai & Lypson 2009).

'Role modeling is at the heart of professional character formation' (Kenny *et al.* 2003, p. 1209). The explicit use of interactions with role models has always been recommended over students entering and exiting the encounters with role models without meeting the true educational potential of such interactions (Kenny *et al.* 2003). Role models can be anyone in the clinical or academic environment (Baernstein *et al.* 2009). However, 'in clinical settings, modeling provides a key opportunity to cultivate professionalism' (Jones *et al.* 2004 p.265) and, as indicated by medical students in the USA in their critical incident narratives, students in clinical rotations mostly learn professional attributes from doctor-patient relationships and teaching from physicians and trainees (Karnieli-Miller *et al.* 2010). Students learn from clinical role models by observing the intended learning goals modeled by role models and / or their unintentional behaviours (Jones *et al.* 2004). Many 'planned' educational activities, however, focused on the latter, as in clinical environments the former is difficult and even unfeasible (Jones *et al.* 2004). Therefore, students may learn both professional and unprofessional attributes which emphasise the need for guided-reflection and facilitated-discussions afterwards (Jones *et al.* 2004). There are several examples of this approach. The final year students of the International Medical University of Malaysia were posted to a rural hospital for clinical training with a special task of observing doctors in those sites in order to understand professional / unprofessional behaviours (Loh & Nalliah 2010). The objective underpinning the task was learning professionalism by observing positive / negative role models. At the end of the

postings, students were able to accurately conceptualise professionalism and reflect upon the practices of their clinical teachers (Loh & Nalliah 2010). The third-year medical students of the University of Minnesota, who undertook the Rural Physician Associate Program (RPAP) in rural communities, used the clinical preceptors as role models for professionalism development and perceived that the nine-month community posting was an ideal environment for developing professionalism (Zink *et al.* 2009). The interaction between clinicians, patients and students during bedside teaching sessions is a powerful opportunity for students to observe professionalism in practice (Monrouxe *et al.* 2009). Clinicians are considered as role models by students during bedside sessions and they learn both negative and positive attributes of doctor-patient relationships (Monrouxe *et al.* 2009). Jones *et al* (2004) at Maryland, USA, used structured assessment of clinical preceptors by students and follow-up discussion, which is the reverse of what usually happens, as a method of teaching professional behaviour to third-year students. The authors argue that this approach encouraged clinical teachers to intentionally model professional behaviours. The students rated this educational experience positively and the teachers that participated in the pilot reported that it had improved the professionalism culture (Jones *et al.* 2004). In the Weill Medical College of Cornell University, the two-week palliative care rotation involves the observation of, and reflection on, role models and professional practices without any patient responsibilities (Fins *et al.* 2003) . This intervention, which was intended to develop ethical, legal, humanistic and communicative qualities, was rated favourably by students.

Not only clinicians, but also basic science teachers are an important source for the development of professionalism among students. If the basic science teaching sessions, such as dissections, are combined with creative projects (e.g. literature, music or media), giving opportunities for students to interact and reflect, the basic science teachers become role models to first-year students to learn the professional attributes of doctors (Shapiro *et al.* 2009), for example, professional behaviour and collegiality (Huggett *et al.* 2008). Both clinical and non-clinical teachers as informal mentors undoubtedly play an important role in fostering professionalism among medical students (Rose *et al.* 2005).

Reflection on practice in fostering professionalism

Students' narratives are a powerful tool for encouraging reflection and understanding hidden aspects of a curriculum, and to assess students' conscientiousness (Baernstein *et al.* 2009; Rees & Monrouxe 2011). In critical incident narratives, the medical students of the Indiana University were expected to reflect upon what they observed in relation to professionalism during their clinical rotations (Karnieli-Miller *et al.* 2010). Analysis of the narratives indicated that students were able to reflect upon several professional attributes of their clinical teachers (e.g. clinicians' relationship with patients and family in clinical setting, and their relationship with students during teaching sessions). Their experiences with both positive and negative behaviours shaped students' professionalism (Karnieli-Miller *et al.* 2010). In a similar attempt, Fard *et al* (2010) identified broader professionalism issues discussed in logbook records of students. Again, the student records have provided students with an opportunity to reflect on what they do. In the USA, the retention of professional attributes such as empathy among students has been

identified with blogging on clerkship experiences and debriefing after significant events (Rosenthal *et al.* 2011).

Workplace events can also be used as a trigger for professionalism education. For example, Cruess and Cruess (2011) argue that collective decisions of doctors such as strike action may influence the professionalism of students and trainees. They suggest that such events should be used to trigger discussions on professionalism, which should ultimately result in students making their own standards in controversial collective decisions. This is also an example where professionalism culture has its influence on individual professionalism.

2.3.1.4. Single versus multi-prong intervention approach

Certain professional attributes, e.g. empathy (Benbassat & Baumal 2004) and ethical behaviour (Feudtner *et al.* 1994), appear to decline gradually as undergraduates progress along the curriculum. Therefore, several authors recommend using a longitudinal multi-prong approach rather than short and one-off interventions in fostering professionalism (Elliott *et al.* 2009; Ginsburg & Lingard 2011; Kumagai & Lypson 2009). Both teachers and students agree that those programmes should commence with the arrival of students at the medical schools (Cruess & Cruess 2006; Krych & Vande Voort 2006). In the USA, a multi-prong approach which included blogging on clerkship experiences, debriefing after significant events, and discussing journal articles, fiction, and film delivered longitudinally has helped retain professional attributes (e.g. empathy) among students (Rosenthal *et al.* 2011). The evaluation carried out by Goldie *et al* at Glasgow Medical School, Scotland,

revealed that a combination of methods (e.g. small group discussions, early clinical exposure, role models) can be used effectively for professionalism education during early undergraduate years, and the integration of professionalism education with the rest of the curriculum was a key expectation of both teachers and students (Goldie *et al.* 2007). A quasi-experimental study conducted in Singapore demonstrated that integration of an ethics course, which included lectures, tutorials and learning from role models, into the formal curriculum imparted more knowledge, built confidence, and made students better professionals compared to *ad hoc* teaching of ethics.

However, the focus of education should be compatible with the different stages medical education, e.g. non-clinical and clinical phases, the phases of observing patient management and engaging in patient management (Ginsburg & Lingard 2011; Hilton & Slotnick 2005). In the design of a professionalism curriculum for pre-clinical students at Keck School of Medicine, California, USA, both didactic methods and interactive group sessions were used (Elliott *et al.* 2009). The designers included more didactic sessions in the first year and more interactive sessions in the second year to suit the development and maturing process of students in the medical school, which was appreciated by the majority of students (Elliott *et al.* 2009). In the Mayo Clinic, USA, professionalism education begins in the first year and is continued towards the final year. Using a similar approach to Keck School of Medicine, didactic lectures, web-based sessions, discussion groups, role-play, simulation using patient-actors, team learning, role-modeling with discussion and reflection have been used primarily for the purpose of shifting emphasis to more interactive methods than didactic methods towards the later years (Mueller 2009).

This does not mean that professionalism education during early years should necessarily be didactic. van Mook *et al* (2010) introduced a longitudinal programme, with plenary lectures, tutorials, and portfolios, to the first three years of undergraduate curriculum. Although students' perceptions about the programme were not reported, the enhanced awareness among students and tutors about professional / unprofessional behaviours due to this programme resulted in an increased number of students with problem behaviours being supported (van Mook *et al.* 2010). However, amidst longitudinal and multi-prong inputs, and individulised feedback on professionalism, the behaviours of a minority of students may still be left unchanged (van Mook *et al.* 2010).

2.3.1.5. Summary

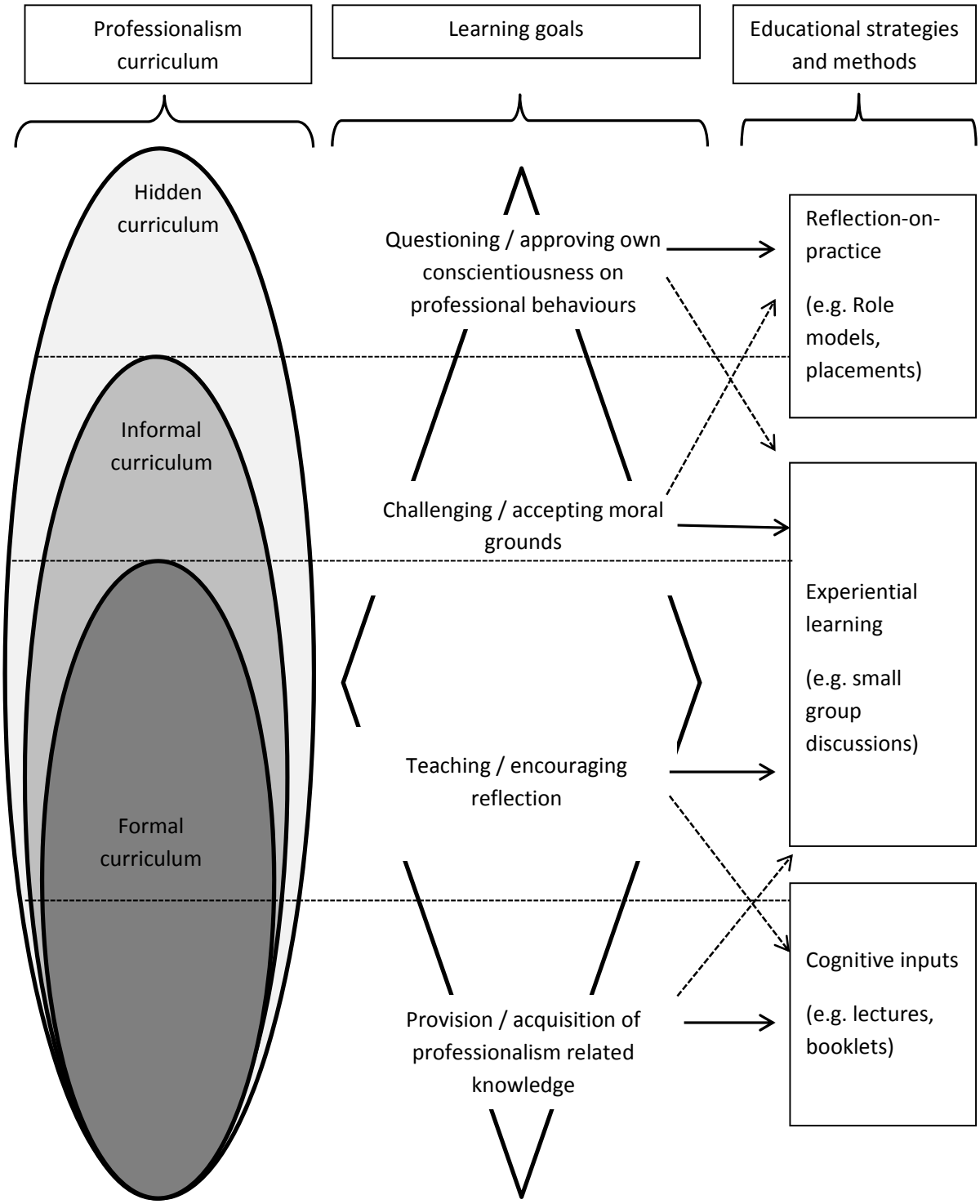
The literature discussed above helps make several meaningful conclusions. Professional attributes can be fostered among students during their undergraduate training. Although informal and hidden curricula may play a greater role than the formal curriculum in this regard, the importance of the latter cannot be ignored.

Medical educationalists have designed dedicated courses or modified existing courses to address professionalism-related outcomes. Although there are reports on one-off interventions directed at a particular stage of a course, many of them are longitudinal, extending across different stages. The latter has been demonstrated to be more effective than the former as professionalism among students tends to erode with time. A gamut of teaching and learning methods has been used in these courses. However, the effectiveness of individual methods or between methods is yet to be determined as the

majority of evaluations are limited to the reactions of participants towards the interventions rather than measuring the actual impact.

In the analysis of teaching and learning strategies, it is evident that there is a hierarchy of learning goals: provision of the knowledge base, encouraging reflection, accepting or challenging moral grounds; and questioning and approving learners' own conscientiousness on professionalism behaviours. These different goals appear to form a taxonomy for fostering professionalism. The provision of cognitive input appears to be a considerable component of the formal curriculum. Encouraging reflection and challenging and accepting moral grounds appear to capture mainly the informal curriculum. Questioning or approving learners' own conscientiousness on professional behaviours may capture the hidden component of the professionalism curriculum. In undergraduate medical education, more emphasis is given to teaching and encouraging reflection and challenging and accepting moral grounds compared to the provision of professionalism-related knowledge and opportunities in the workplace to question or approve students' own conscientiousness on professional behaviours. Provision of knowledge is achieved primarily through cognitive inputs and to some extent through experiential learning. Reflection is encouraged mainly by experiential learning with some contribution from cognitive inputs. The goal of challenging or accepting moral grounds is achieved primarily by experiential learning and to some extent through reflection-on-practice. Questioning or approving own conscientiousness is achieved primarily through reflection-on-practice with some contribution from experiential learning. (Figure 2)

Figure 2 – Relationship between professionalism curriculum, learning goals and educational strategies as evident in the literature



Teaching / learning approaches targeted at the lower levels of the taxonomy (e.g. providing necessary knowledge) contribute to achieving the higher levels (e.g. questioning or approving learners' conscientiousness on professional behaviours). However, targeting the lower levels alone would not necessarily be effective as higher level approaches appear to be more effective in fostering professionalism among students. Allocating time in the formal curriculum for educational methods which could address the higher learning goals will help encroach into informal and hidden curricula. For example, if time is allocated in clinical rotations for discussing the professionalism dilemmas encountered by students during their day-to-day work it will be an opportunity for them to reflect on informal and hidden aspects and learn from their own experience.

Similar conclusions have been made by O'Sullivan *et al* (2012) in their recently published review of literature on professionalism education.

This section of the literature review informed the project in several ways.

- The review helped in understanding of the formal curriculum of professionalism and its intentions, which is the first step to understanding the informal and hidden components of a curriculum.
- It reiterates the impact of institutional culture on fostering professionalism, of which the formal curriculum is part.
- Professionalism in any individual is a state that can be reformed to some degree; therefore, understating and manipulation of the institutional professionalism culture would be useful in facilitating the reform process.

2.3.2. How should professionalism be assessed?

Assessment of professionalism has focused on assessing individuals and groups or institutions. This section focuses on the assessment of professionalism of individuals at undergraduate level, and the assessment of institutional professionalism (professionalism culture) is in a subsequent section. As with any student assessment, gaining insight into why, what, how, who, where and when the professionalism assessments should be carried out is useful (Hawkins *et al.* 2009).

In the literature search key words used were primarily focused on but not confined to: medical, professionalism, professional behaviour, professional attitudes, undergraduate, assessing, evaluating, monitoring.

2.3.2.1. Why should professionalism be assessed?

Until recently, assessment of individuals in medical education has focused primarily on medical knowledge, practical skills and clinical skills and, professionalism has been an opportunistic and implicit component of such assessments with a minimal impact on pass/fail decisions (Arnold 2002; Ginsburg *et al.* 2000). The main reason for this was that fitness to practise medicine was defined as physical and mental fitness alongside clinical competence (Parker 2006). However, more recently, the professional conduct of doctors has been subjected to increased public scrutiny in addition to their knowledge or technical skills (Papadakis *et al.* 1999). Furthermore, with enhanced emphasis on the social accountability of medicine, the medical profession is expected to assure its clientele that its members are both ‘safe’ and ‘professional’ practitioners (Boelen & Woollard 2009). As

a result, today, fitness to practise medicine is defined broadly, encompassing not only clinical competence and lack of impairments, but also professionalism (Parker 2006). Educationally, it is widely acknowledge that 'assessment drives learning', especially at undergraduate and postgraduate levels (Epstein 2007). Therefore, professionalism should be assessed to encourage learning of professionalism (Cruess & Cruess 2006). Thus, the assessment of professionalism in medicine has become a formal, moral and educational obligation for medical schools.

Although the purpose of every assessment is the provision of feedback for further learning, traditionally, the purpose of an assessment is dichotomously defined as to: provide feedback (formative), and make pass / fail decisions (summative) (Epstein 2007). For centuries, there has been more emphasis on the latter, i.e. assessment *of* learning (Lynch *et al.* 2004). This particularly has heightened the challenge in assessing professionalism; both educators and students alike perceiving that, unlike scientific facts, for the issues of professionalism, there are no 'right' or 'wrong' answers (Ginsburg *et al.* 2000; Parker 2006). This argument was counteracted to some extent with the introduction of community standards, professional codes and legislations, and using them as benchmarks (Parker 2006). However, there is a growing emphasis towards the 'assessment *for* learning', i.e. the main purpose of an assessment is to provide feedback for future learning and development (Lynch *et al.* 2004). Although this approach has its own challenges, it has more acceptability and impact than the summative approach in the assessment of professionalism (Ginsburg *et al.* 2000; Lynch *et al.* 2004; van Mook *et al.* 2010). However challenging it is, criterion-based summative assessments are essential to

ensure the social accountability of the medical profession (van Mook *et al.* 2010). On the other hand, students should be given meaningful feedback on their professional behaviours to facilitate their professionalism development (Lynch *et al.* 2004; van Mook *et al.* 2010), which has been demonstrated to be occurring throughout their training (Hafferty 1998). Therefore, a combined and balanced approach, which includes both formative and summative assessments, may be the way forward (van Mook *et al.* 2010).

2.3.2.2. What should be assessed under 'professionalism'?

It is relatively easy to measure students' awareness of what is expected by the profession (e.g. using written tests) and their ability to recongise or demonstrate professional behaviours in simulated environments (e.g. using the OSCE) (Parker 2006; van Mook *et al.* 2009a). Although important and contributory, these measurements alone do not serve the ultimate goal of professionalism assessment, which is the determination of whether students understand how to behave professionally, i.e. whether they act with professional diligence (Rees & Knight 2007). The assessment of the latter is challenging and the obscurity of the challenge is aggravated in the context of summative assessments (Parker 2006; van Mook *et al.* 2009a). Despite the challenge, the majority of professionalism assessments attempt to assess the actual professional behaviours of students (Parker 2006; van Mook *et al.* 2009a; Wilkinson *et al.* 2009).

The debate on whether to measure the professional attitudes, professional behaviour or combination of both is ongoing. In the behaviour-centred approach to assessment, professionalism requires to be defined as a set of abstract attributes as the first step

(Arnold 2002; Parker 2006). Most definitions of professionalism, however, are broad and not spelt out in terms of specific attributes, making this step problematic for many assessment setters (Wilkinson *et al.* 2009). Even if the attributes are identified, they may not be interpreted accurately by the raters unless they have been operationalised in behavioural terms (Mazor *et al.* 2008). The assessors may need to use multiple assessment tools to encompass all professional bahaviours, as a single tool may not be able to reliably capture them all (Wilkinson *et al.* 2009). The 'behavioural approach' to professionalism assessment has been subjected to criticism as, especially in assessment situations, professional attitudes may not accurately be reflected in students' behaviours; they may 'demonstrate' the expected behaviours rather than the actual manifestations of their attitudes (Rees & Knight 2007). As a result, more and more assessment systems attempt to encompass both professional attitudes and behaviours. In circumstances under which this is not always possible, it is recommended to gather contextual information to substantiate the behaviours observed (Ginsburg *et al.* 2004; Rees & Knight 2007).

2.3.2.3. How and where should professional attitudes and behaviours be assessed?

Professionalism of students has been assessed by using written assessments in classroom or examination hall environments, and examination or observations in simulated or workplace-based environments.

a. Written assessments

The assessment of professionalism-related knowledge in examination-hall environments should be part of the assessment process (Wilkinson *et al.* 2009). For example, context-

rich multiple choice questions may be used in the assessment of ethical principles and medico-legal aspects of professionalism (Kao 2006). However, most commonly used are the scenario-based written tests, in which students are expected to propose with reasons the most appropriate approach to dealing with professional dilemmas (Wilkinson *et al.* 2009). These tests have the potential to assess higher order thinking such as moral reasoning and ethical decision making (Wilkinson *et al.* 2009). In the pen-and-paper format of this test, the inter-assessor reliability among the faculty members may be poor (Ginsburg *et al.* 2009). With the advancement of technology, these tests have evolved from pen-and-paper to video and follow-up interviews, which has necessarily helped improve the authenticity of assessments and exploration of the thinking behind responses (Ginsburg *et al.* 2007). However, given the nature of the construct, the assessment of knowledge alone will neither be adequate nor entirely appropriate in the assessment of professionalism (Wilkinson *et al.* 2009).

b. Simulation-based assessments

The Objective Structured Clinical Examination (OSCE) and simulated ward exercises are examples of methods used in this category. Especially in the undergraduate settings, OSCE stations have been used to assess professionalism with stations dedicated entirely to the purpose or with stations which assess professionalism together with other competencies (e.g. clinical skills) (Ginsburg *et al.* 2000). These stations frequently use standardised /simulated patients or artificial models as the 'assessment material' (Ginsburg *et al.* 2000; Wilkinson *et al.* 2009). The OSCE has been used alone for the purpose of assessing professionalism in the Emory University School of Medicine, Atlanta, by Wallenstein *et al*

(2010); they claim that the method is valid and reliable. The ability to bring the context into the assessment scenario (e.g. time pressure, interaction with the scenario) is a key advantage of the OSCE compared to written examinations. However, there are several issues with using the OSCE for assessing professionalism. The artificiality of OSCE situations has posed more challenges to the assessment of professionalism than that of clinical / practical skills (Ginsburg *et al.* 2000). Unlike OSCE marks for medical knowledge and clinical skills, OSCE marks for professionalism may not predict professional behaviour of students in the workplace (Wallenstein *et al.* 2010). When students are aware that they are being assessed for their professionalism, they may fake their attitudes and behaviours, especially in a summative assessment (Rees & Knight 2007). Regardless of who the assessors are (clinicians or standardised / simulated patients), the reports on the reliability of the OSCE in the assessment of professionalism are still inconclusive (Ginsburg *et al.* 2000). Delineating professionalism or its components into its sub-constituents and the professional expectations into a rating scale, a process which is important for a valid and reliable OSCE, may cause ambiguity among the assessors and lead to non-discrimination of students (Ginsburg *et al.* 2000). For example, an OSCE has not helped discriminate between students' communication skills and their ability to analyse a given problem on ethical grounds; the good communicators scored highly for ethical analysis despite their analyses being poor (Arnold & Forrow 1993). Similar observations have been made when standardised patients were used as assessors (Prislin *et al.* 2001; van Zanten *et al.* 2005). The use of simulated ward exercises with high fidelity simulators helped improve the authenticity of the assessment experience in assessing professionalism (e.g.

prioritisation, team-work) of students in a near-realistic setting (Ker *et al.* 2006). The simulation-based assessments would, however, be more effective if they are supplemented with a post-encounter probe of students' thinking behind the responses or actions (Lynch *et al.* 2004).

c. Work-place based assessments

Given the nature of the construct, work-place based assessments are the preferred method of assessing professionalism (Wilkinson *et al.* 2009). It appears that, in clinical settings, students have been assessed by examining their ability to reflect on practice and by observing their behaviours.

- Examination of reflection-on-practice

The ability to reflect is the basis of professionalism (Fryer-Edwards *et al.* 2006; Wilkinson *et al.* 2009). Self-reflective ability of students, trainees and doctors on their own practices, therefore, has been a focus of assessment of professionalism, and it has been assessed primarily through portfolios and critical incident reports (Wilkinson *et al.* 2009). Critical incident reports are short, written narrative accounts based on incidents, which have taken place in day-to-day practice (Branch 2005). The authenticity of the incident and the depth of reflection are the key features of such accounts (Branch 2005). Portfolios, on the other hand, are a collection of evidence with accounts of self-reflection and have been used for the assessment of professionalism of undergraduate medical students (Davis *et al.* 2001). In critical incident reports and portfolios, the authenticity and validity of assessment, and the educational impact is high as they are based on routine practice

(Branch 2005; Davis *et al.* 2001). However, they may become questionable in terms of reliability (Driessen *et al.* 2005). In addition, the self-assessment skills of students are unreliable (Epstein & Hundert 2002), and self-assessments, especially as a part of summative assessments, may be unacceptable to students as well as teachers (Rees & Shepherd 2005b).

- Observations

Supervisor and peer observations have been used alone for the assessment of professionalism of students in clinical settings. Originally, the Mini Clinical Evaluation Exercise (mini-CEX) was designed to assess mainly clinical skills of students and trainees reliably by their supervisors during routine practice (Norcini *et al.* 2003). However, professionalism related to the encounter under observation has been a part of the mini-CEX overall assessment (Norcini *et al.* 2003; Wilkinson *et al.* 2009). The Professionalism Mini Evaluation Exercise (P-MEX), an adaptation of mini-CEX specifically to assess doctor-patient relationship skills, reflective skills, time management, and inter-professional relationship skills of students, collates the judgements of supervisors in different clinical settings (e.g. wards, clinics) (Cruess *et al.* 2006). It has been reported as a measure with adequate reliability, acceptability to both students and supervisors, and feasibility to be used in clinical contexts (Cruess *et al.* 2006).

Peer assessment, on the other hand, has been demonstrated to be an acceptable and effective method of assessing professionalism of undergraduates, at least in the USA (Arnold *et al.* 2007; Nofziger *et al.* 2010), provided that the assessment is anonymous,

focused on both professional and unprofessional behaviours, used for formative purposes, used to reward exemplary behaviours and used to support repeated lapses of professionalism (Arnold *et al.* 2007). In the UK, peer assessment of professionalism has been positively correlated with professionalism as measured by routine behaviours of students (e.g. attendance, the on-time submission of assignments) (Finn *et al.* 2009). Speyer *et al* (2011) in their review claim that the majority of the peer assessment tools for professional behaviours are context-specific and lack validity and reliability, thus limiting their wider use.

The assessment of professional behaviours of students by patients and nurses has been recommended (Wass *et al.* 2001). However, it has rarely been used as the only method to assess students in the workplace environment (Veloski *et al.* 2005; Wilkinson *et al.* 2009). This may be due to the fact that student-patient or student-nurse interactions in clinical settings may not be sufficient to provide a reliable and all-encompassing judgement on the professional behaviour of every student. The judgements of patients and nurses, however, are useful components of multi-source feedback (Veloski *et al.* 2005; Wilkinson *et al.* 2009). Multi-source feedback, i.e. 360 degree assessment, collates judgments from different members of the healthcare delivery team (e.g. supervisors, peers, nurses, patients) on professional behaviour of students (Shrank *et al.* 2004; Wilkinson *et al.* 2009). This method has been commended as a comprehensive assessment of professionalism as a single candidate is evaluated multiple times by multiple raters in multiple settings (Wilkinson *et al.* 2009). It has been demonstrated that multi-source feedback is valid, reliable (at least in the assessment of doctors) (Whitehouse *et al.* 2005) and acceptable

(Rees & Shepherd 2005a) in the assessment of professionalism. However, variation in contact time and lack of discrimination by assessors may limit the benefits of multi-source feedback especially when they are used to assess undergraduates (Rees & Shepherd 2005a). Therefore, multi-source feedback may be more suitable at postgraduate level (Lynch *et al.* 2004).

2.3.2.4. Who should assess professional attitudes and behaviours in simulated environments and workplaces?

In assessing professionalism of students in simulated environments, the faculty and physicians (Ker *et al.* 2006) and / or simulated patients (Ginsburg *et al.* 2000) are very effective resources. However, the inter-rater reliability of these raters may be problematic especially in the absence of a clear definition of the construct being assessed (Ginsburg *et al.* 2000).

The professionalism of medical undergraduates in the workplace can be assessed either by the faculty / physicians, or via multiples sources (Veloski *et al.* 2005). In workplace settings, the faculty, peers and patients as assessors identify unprofessional behaviours of students, and agree upon the unprofessional nature of such behaviours more readily than that for professional behaviours (Wilkinson *et al.* 2009). Trained faculty and physicians reliably judge the professionalism of students and trainees (Ginsburg *et al.* 2000). However, they are usually reluctant to put unprofessional behaviours on record making the discrimination between professional and unprofessional student / trainees difficult (Frohna & Stern 2005; Ginsburg *et al.* 2000). Given the fact that peers have more

frequent, close and varied contact, they may be better than the faculty and physicians in assessing interpersonal skills (Ginsburg *et al.* 2000). However, in the context of summative assessment, their rating may be or may not be discriminating (Ginsburg *et al.* 2000). In the assessment of professionalism of postgraduate trainees and practitioners, nurses have mirrored patients' judgements (Boon & Turner 2004). However, the judgements of nurses and patients may be influenced by multiple factors unrelated to professional attitude or behaviours (e.g. 'good' relationship or continuity of care) which may ultimately lead to lack of reliability (Ginsburg *et al.* 2000).

2.3.2.5. When should professionalism be assessed?

The assessment of professionalism should commence in early undergraduate years and continue longitudinally to physician stage through the postgraduate training stage (Hodges *et al.* 2011; Lynch *et al.* 2004). The orientation of assessment, however, may have to be tailored to suit the level of moral development and professional training (Baldwin & Self 2006; Hilton & Slotnick 2005). Professionalism should be assessed frequently at every available opportunity, but with adequate space for the assessee to receive feedback and to improve (Hodges *et al.* 2011). In early undergraduate years, the focus may be on professionalism during academic activities (Bryan *et al.* 2005; Roff *et al.* 2011). With the exposure of students to the clinical environment in later years, professionalism should encompass the new opportunities and challenges in clinical environments (Roff & Dherwani 2011).

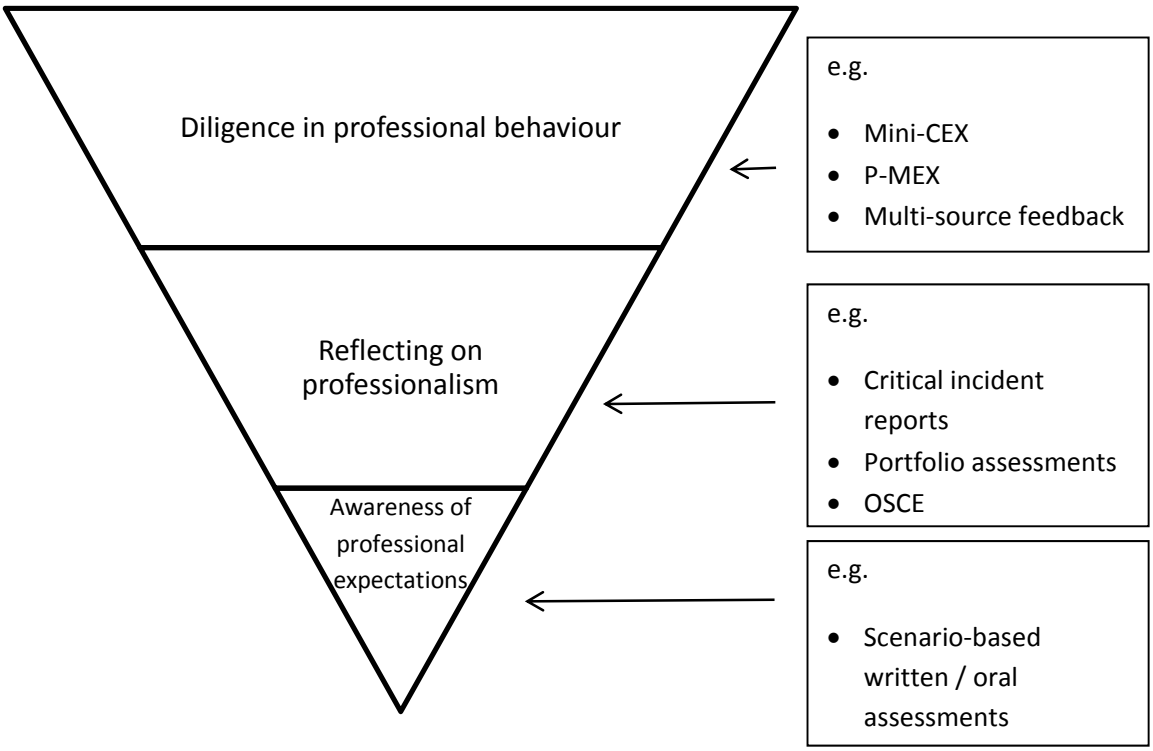
2.3.2.5. Summary

Professionalism should be assessed formatively to facilitate the development of professionalism in the student, and summatively to ensure fitness to practise. Professionalism assessment should focus more on attitudes and behaviours than students' knowledge of professionalism. The former can be assessed by examination of students' ability to reflect-on-practice, and observation of students' behaviour in simulated and workplace environments. Variations in self-assessment ability, however, may adversely affect the assessment of reflection on practice. In the workplace, to achieve acceptable reliability of summative assessments, professionalism assessments may target the observable professional behaviours, but underlying attitudes and contextual information should also be considered. Workplace assessments should ideally be multisource. However, faculty / physicians and peers may be better assessors of students' professionalism than nurses and patients. The assessment of professionalism should commence in the early years of undergraduate education and should be continued longitudinally. However, the assessments should be tailored to the moral development of students and their level of training in order to be more effective.

It appears according to the literature that assessment of professionalism especially for summative purposes fits into a hierarchy of assessments: the awareness of professional attitudes and behaviours expected; the ability to reflect on the professional aspects in routine practice; and the diligence of professional behaviour in the workplace. For obvious reasons, assessment setters place more emphasis on the assessment of diligence in

professional behaviour compared to the awareness of professional expectations. (Figure 3)

Figure 3 – The hierarchy of professionalism assessment and methods used at each level



The current project was informed in this section by the literature in several ways, which are summarised below:

- The influence of the professionalism culture on the professionalism of individuals may be evident especially in the assessment of professionalism in the workplace.
- Understanding, the context is necessary to accurately interpret the assessed professionalism in individuals. Therefore, understanding professionalism culture of the institution is important.

- Both attitudes and behaviours should be focused on in the measurement of professionalism
- The behaviours and attitudes chosen for measurements, however, should be appropriate to the stages of moral development and training.
- Comparatively, faculty and peers are deemed to be reliable in making judgments of professionalism
- Professionalism assessment itself may be a cause for changes in professional behaviour.

2.4. Exploring institutional professionalism culture as a part of undergraduate professionalism education

This section focuses on understanding institutional culture and its constituents, how it has been measured historically, and identifying a theoretical framework for measuring it.

2.4.1. Institutional culture

In the literature search key words used were primarily focused on but not confined to: culture, organisational, institutional, measure, evaluation.

'Culture' is defined by the Merriam-Webster online dictionary as 'the integrated pattern of human knowledge, belief, and behaviour that depends upon the capacity for learning and transmitting knowledge to succeeding generations'. 'Organisational culture', which is basically a qualification of the generic definition of culture focusing more on work, working relationships and the working environment (Schein 1984), is a relatively new concept in the fields of business and management. Although there are several definitions

for ‘organisational culture’ the relevance and applicability of a particular definition to a given context depends very much upon the nature of the institution and its inhabitants (Scott *et al.* 2003a; Scott *et al.* 2003b). Where healthcare professionals and institutions are concerned, it is more appropriate and relevant to conceptualise organisational culture as an anthropological metaphor, i.e. the organisation of social relations among humans (Hudelson 2004; Scott *et al.* 2003a). Accordingly, the term organisational culture in the context of healthcare institutions can be described as ‘ a shared set of (implicit and explicit) values, ideas, concepts and rules of behaviour that allow a social group to function and perpetuate itself’ (Hudelson 2004). The behaviour of the members is influenced and guided by the culture of the organisation they belong to, and the pattern of behaviours of members often reflects the features of the organisation’s culture (Hudelson 2004). The behaviour of members together with their attitudes and beliefs, therefore, is a useful and recognised parameter in measuring or describing the culture of an organisation or institution. The culture of a social group, e.g. culture within an institution, however, is not static but dynamic, and is responsive to internal and external influences (Hudelson 2004).

As the definition suggests, the measurement of an organisation’s culture needs a blended approach, which takes in to account the attitudes, beliefs and behaviours of the people and internal and external factors influencing them. Hudelson (2004) prefers the traditional qualitative approaches in investigating organisational culture, especially in comparing the organisational culture and the quality of the institution’s outcome, as behaviours are more explainable than quantifiable. This notion was successfully implemented by Zilwa (2007) in

examining the impact of the organisational culture on the adaptation of academic units to change in Australian universities. However, Scott *et al* (2003a), in their review of instruments for measuring organisational culture quantitatively, identified 13 instruments with acceptable validity and reliability, and proposed four of them for use in contexts of healthcare institutions. Although the ability of these inventories to establish any relationship between organisational culture and the performance outcomes of the institution is inconclusive, they have been shown to be successful in describing the culture (Scott *et al.* 2003b). The content of the instruments identified by Scott *et al* (2003a) (Table 4), however, does not represent what is meant by professionalism in medicine (Quaintance *et al.* 2008). In medicine, it appears that the prefixes 'organisational' and 'professional' provide two different perspectives of 'institutional culture'. The evidence regarding organisational culture supports the notion that it can be quantitatively measured and explained which may also be applicable to the measurement of institutional professionalism culture.

Table 4 - Examples of content included in measures of organisational culture (Scott *et al.* 2003a)

Leadership style
Bonding systems
Prioritisation of goals
Character of organisation
Managers' style
Cohesion
Rewards
Expectations that guide thinking
Behaviour of group members

2.4.2. Measures of professionalism culture / environment of institutions training medical undergraduates

In the literature search, key words used were primarily focused on, but not confined to; medical, professionalism, professional behaviour, professional attitudes, culture, climate, environment, inventory, survey, scale, measure. These, however, largely yielded unrelated results. This may be due to the considerable variation in the nomenclature of such instruments in the literature. Therefore, a search for individual articles cited in reviews was carried out using Google scholar.

Medical educationalists and researchers have developed several quantitative measures of the professionalism environment and culture in the context of medical education. However, some of these focused on assessing the professionalism environment of postgraduate (Cohen & Patten 2005; Reid *et al.* 2006) and continuing medical education (Marshall 1999; Passmore & Leung 2002) settings, and some focused on undergraduate settings which used qualitative approaches (Karnieli-Miller *et al.* 2010; Lempp & Seale 2004). As the primary aim of this project was to quantifiably assess the professionalism culture in the undergraduate setting, the quantitative measures which focused on undergraduate settings or discussed an undergraduate component were included.

The quantitative measures identified have been referred to variously as ‘scales’, ‘inventories’ and ‘surveys’ by their inventors. For the purpose of this discussion all these will be referred to as ‘measures’**.** Twelve measures, which emerged in the literature during the last 15 years (the new era of professionalism), are discussed based on the

popularity among other researchers, variety in methodology and similarity to this project. Below, the key features of each of these measures are described. These features are critically analysed and compared in the subsequent discussion.

2.4.2.1. Measures of professionalism culture based on a predetermined professionalism framework

Several measures have been used in the USA to measure the professionalism environment based on the ABIM framework of professionalism (Table 2, p.1).

a. ABIM Scale to Measure Professional Attitudes and Behaviours in Medical Education (Arnold *et al.* 1998)

In 1998, Arnold *et al* developed this measure to assess the professionalism environment in US medical training institutions, using the perceptions of students and residents. The measure consisted of 14 items developed by the authors themselves to reflect the domains of professionalism as defined by the ABIM and the behaviours of consultants and residents. The respondents were expected to indicate the frequency of each behaviour ('never' to 'always') in a nine-point scale based on what they have observed during the clinical rotations.

A set of 757 questionnaires was distributed by post with 529 returns (response rate = 69%), which were included in the analysis. The authors excluded two items from the analysis due to missing data and poor correlation between items and the total score. An exploratory factor analysis of the remaining 12 items yielded three principal components, which corresponded with the ABIM framework (five items under 'excellence', four items

under 'honesty / integrity', and three items under 'altruism / respect'). The internal consistencies of both instrument (Cronbach alpha = 0.71) and domain (Cronbach alpha = between 0.72 – 0.59) levels appear to be acceptable. Although the total and domain scores for each participating institution have been computed, only the total score and the score for the domain of 'excellence' have demonstrated discrimination ability between institutions. (Table 5)

Table 5 – Examples of items included in the ABIM Scale to Measure Professional Attitudes and Behaviours in Medical Education (Arnold *et al.* 1998)

ABIM domain	Example items
Excellence	During this residency training, I have met individuals whom I consider role models
	I have been instructed to withhold data from a patient's chart without being given an explanation from my senior resident or attending physician.
Honour / Integrity	I have observed my resident colleagues lie to a patient.
	I have been urged by my resident colleagues to copy their history and physical exam rather than gathering my own information from the patient.
Altruism/ Respect	I have observed my resident colleagues making derogatory statements about other medical/surgical specialists or specialty groups, subspecialists or subspecialty groups, or other health care professionals.
	I have observed resident colleagues scheduling tests or performing procedures at times that are more convenient for themselves than for the patient.

b. Pennsylvania State University College of Medicine (PSCOM) Professionalism Questionnaire (Blackall *et al.* 2007)

This questionnaire was developed to evaluate professionalism at institutional level and determine the effects of socialisation on its development. In the development of items the authors used the ABIM professionalism framework as the basis, and representative items were initially identified through a literature review. The number of items was subsequently reduced to 36 (six items for each ABIM domain) through a three-round

Delphi process. The six-member Delphi panel included academics involved in professionalism research. The PSCOM Professionalism Questionnaire was made into four parallel forms to suit different groups in the medical school; medical students, residents, basic science faculty and clinical science faculty, by making minor changes to the items. The five-point rating scale required the respondents to rate the extent to which each item represents their own definition of professionalism (from 'never' to 'great deal'). In addition, the respondents were asked to rank order the representativeness of the six items included under each domain.

The self-administered questionnaires were returned by 765 students and staff members of the medical school (response rate = 51%). A principal component analysis, carried out to test the hypothesis that the clustered items represent the domains, generated seven factors instead of the six domains expected. These factors were renamed as; accountability, enrichment, equity, honor and integrity, duty, and respect. The examples of the items clustered under the new domains are included in Table 6.

Table 6 - Domains and examples of items in the PSCOM Professionalism Questionnaire (Blackall *et al.* 2007)

Domains	Sample items
Accountability	Works collaboratively and respectfully within a team to the benefit of improved patient care or to contribution of research
Enrichment	Shows a willingness to initiate and offer assistance towards a colleague's professional and personal development
Equity	Adopts uniform and equitable standards for patient care
Honour and integrity	Upholds scientific merit and bases decisions on scientific evidence and experience
Altruism	Shows compassion
Duty	Participates in correlative action processes towards those who fail to meet professional standards of conduct
Respect	Avoids offensive speech, unkind comments and unfair criticisms to others

c. Professionalism Climate Instrument (Quaintance *et al.* 2008)

The purpose of this instrument was to measure the professional behaviours of undergraduate students, postgraduate trainees and faculty (professional behaviour assessment) and the extent to which professionalism teaching takes place in the clinical environment (professionalism teaching behaviour assessment) in the University of Missouri–Kansas City School of Medicine, USA. The component on professional behaviour assessment consisted of 12 items (Table 7); each item was rated by students in relation to colleagues, residents and the faculty on a four-point frequency scale ('observed mostly' to 'observed rarely'). These items were extracted from what students expressed in focus group discussions which authors believed were representative of the ABIM professionalism domains. The component of professionalism teaching assessment consisted of 10 items focused on professionalism of the faculty related to teaching and these were identified by reviewing the literature on the clinical teacher as a role model (Table 7). This component was developed in two versions: (a) the student version possesses the same rating scale as the behaviour assessment component and (b) the faculty version, which is a self-assessment of their own teaching practices, and possesses a four-point rating scale of 'usually done' to 'never done'.

Table 7 - Components and sample items of Professionalism Climate Instrument (Quaintance *et al.* 2008)

Component	Examples of items
Professionalism behaviour assessment	show disrespect to patients, students, faculty staff or other healthcare personnel
	Make selves look good at the expense of others
	Complain about professional obligations
Professionalism teaching assessment	Act professionally in relation to patients, students, colleagues and staff
	Teaches about professionalism
	Set clear expectations for students' professional behaviour

In the pilot study, the instrument, administered face-to-face, was responded to by 242 students (response rate = 68%; 128 pre-clinical) and 28 faculty members of the same institution. In the analysis of results, all five subscales (i.e. professionalism behaviours of: (1) students, (2) residents, and (3) the faculty, and the assessment of teaching practices by (4) students and (5) the faculty themselves) demonstrated high internal consistency (Cronbach alpha > 0.75). The findings of this study were compatible with other studies that claimed the higher the extent and depth of professionalism teaching the better the professionalism environment (Monrouxe *et al.* 2011). Pre-clinical students, in general, tended to rate professionalism behaviours of their peers and teachers more positively than clinical students (Patenaude *et al.* 2003; Satterwhite *et al.* 1998). This instrument was used subsequently to study the correlation between burnout of students and empathy, and the professionalism culture of the institution, which revealed that the high scores for professionalism culture correlate with high scores for empathy and low scores for burnout (Brazeau *et al.* 2010). This may lead to two conclusions: the better the professionalism culture the more the students become empathic and there is less chance of burnout; or

the students who are empathic and enjoy the course rate the professionalism culture more highly.

2.4.2.2. Instruments measuring professionalism culture/environment not based on a pre-determined framework

a. Medical Professionalism Attitudes and Beliefs Questionnaire (Kalet & Steven 2004)

The purpose of this paper-and-pencil questionnaire was to determine medical students' feelings and beliefs, and the impact of educational interventions on professionalism in a medical school in the USA. It has 44 items selected by the authors from answers given to the question, 'What is medical professionalism?' in focus group discussions with students, and from the literature. Based on the responses of 235 students (first to fifth year), a factor analysis was conducted which yielded four constructs; (i) professional behaviour in the medical school, (ii) knowledge of the definition of professionalism, (iii) confidence in one's professionalism, and (iv) the importance of duty and social justice. Both the overall and domain-level reliabilities were reported to be high (Cronback alpha >0.73). This report, however, does not include examples of the items, a description about the type of rating scale used and a discussion on the four domains identified.

b. Web-based Moral Distress Survey (Wiggleton *et al.* 2010)

This online survey was developed in Vanderbilt University School of Medicine, USA, to assess the moral distress among students resulting from unprofessional events happening in clinical environments. It consists of 55 events, which may cause moral distress among medical students and was adapted from a similar scale used in the field of nursing, and

supplemented by a literature review. The events were validated by a small group of academics and students before finalisation. The rating scale had two components: (i) in relation to each event, students were expected to indicate the frequency at which they observed such an event (a five-point scale from 'never' to 'very frequently'); and (ii) the extent of distress caused (a three-point scale 'no distress' to 'severe distress'). Examples of the 'events' used are included in (Table 8).

Table 8 - Examples of 'events' included in the Moral Distress Survey (Wiggleton *et al.* 2010)

Members of my team "bad-mouthed" other services
A member of my team made disparaging or demeaning remarks about one of our patients
Poor communication between multiple teams that were collaborating in the care of one my patients negatively affected his or her care
Our team continued life support, even though I thought it was not what the patient would have wanted
The attending physician or resident answered a patient's questions inadequately or simply ignored them

This online instrument, sent by email to fourth-year students of Vanderbilt University School of Medicine, was responded to by 60% (n = 64) of the cohort. The analysis of the results revealed: a high internal consistency (Cronbach alpha = 0.95); at least half of the morally distressing events included in the survey had occurred in the clinical environment; and mild to moderate distress among students by 19 of these events. The authors calculated individual moral scores for each student, which correlated well with the frequency of events observed by the student concerned. The morality score of the cohort and number of incidents reported could be used to explain the moral outlook of the institution.

c. Jefferson Scale of Lifelong learning (Hojat *et al.* 2003b; Hojat *et al.* 2009a; Wetzel *et al.* 2010)

This scale was developed to measure the orientation of physicians and students towards lifelong learning in an institutional environment and exists currently in two separate versions; physicians and students. Hojat *et al* (2003b), from the Jefferson School of Medicine, USA, originally developed the physician version to measure the attitudes of the clinical teachers towards lifelong learning in the USA, which is an important aspect of professionalism globally (GMC 2006; Jotkowitz *et al.* 2004; Swick 2000). The authors used a two-step process. In the initial step, 40 features of lifelong learning were identified from the literature by a group of 12 members of staff involved in the project. In the next step, these items were condensed to 37 through a Delphi process with the participation of 28 members of the academic staff (excluding the authors). Each item was scored on a four-point Likert-type agreement scale ('strongly agree' to 'strongly disagree'). In addition, the respondents were asked to rate themselves on a 10-point global rating scale (1 = not committed to lifelong learning at all, 10 = a tireless advocate of lifelong learning). This global rating scale was used as a criterion measure for the validity study (Hojat *et al.* 2003b). The scale was piloted as a postal questionnaire among one third of the total clinical staff of the Jefferson University Healthcare System, selected randomly, and yielded 160 responses (response rate = 43%). In the analysis, the authors included the items, which demonstrated a statistically significant correlation (r>0.2 and p <0.05) with the global score and a rotated factor loading coefficient greater than 0.40; only 19 items met this criteria (Hojat *et al.* 2003b). In the exploratory factor analysis, five factors (need

recognition, research endeavour, self-initiation, technical / computer skills and personal motivation) were identified as domains of lifelong learning, which were demonstrated to be conceptually relevant to the features of lifelong learning as described in the literature (Hojat *et al.* 2003b). The authors claim that this concordance provides supportive evidence for the construct validity of the scale. Further validity evidence was provided by using the 'method of contrasting groups' (Hojat *et al.* 2003b quoting Anastasi, 1976), where means scores for males and females for each item have been compared with the *t*-test and effect size. The internal consistency, calculated by means of Cronbach alpha, indicated that the scale has acceptable internal reliabilities at both scale (0.93) and domain levels (0.99 – 0.64). The means scores for the scale, calculated by averaging the responses of all respondents were used to interpret the tendency for lifelong learning in the given environment.

This scale has been validated further among a larger number of physicians outside the university setting in the USA (Hojat *et al.* 2009a). To ensure the face validity of the 19 items, the authors obtained feedback from 30 such physicians and they suggested the exclusion of five items, leaving 14 items in the final scale (Hojat *et al.* 2009a) (Table 9). The revised scale was field-tested with 3195 physicians outside the university system, who represented 60% of the graduates of Jefferson medical school. To establish the test-retest reliability, four months after the initial launch, the scale was posted again to 200 respondents selected randomly of which 137 (64%) responded. Instead of the five factors observed in the first study, a principal component analysis of the responses generated only three factors: learning needs and motivation, attention to learning opportunities, and

technical skills in seeking information. As in the first study, the internal consistencies of revised the items at both scale and domain levels were acceptable (Cronbach alpha > 0.80). The test-retest reliability was reported to be acceptable (0.72 – 0.77). As there were only 14 items in the latest version, scores of lifelong learning orientation varied between 14 and 56. In addition to its primary purpose, the scale was able to identify the differences between physician groups, e.g. academic and non-academic physicians (physicians involved in teaching and research activities in addition to patient care had better orientation towards lifelong learning than physicians involved only in patient care), and academic achievers and non-achievers (physicians with postgraduate academic qualification or ones who have done well in the medical school were better oriented to lifelong learning than the others) (Hojat *et al.* 2009a) .

Table 9 - Examples of items included in the Jefferson Scale of Physicians' Lifelong Learning (Hojat *et al.* 2009a)

Searching for answers to questions is, in and by itself rewarding
I enjoy reading articles in which issues of my professional interest are discussed
One of the important goals of medical school is to develop students' lifelong learning skills
I recognise my need to constantly acquire new professional knowledge
My professional approach in finding an answer to a question is to search the appropriate computer databases

In 2010, Wetzel *et al* validated the Jefferson Scale of Lifelong Learning for Physicians for the use of students in the Virginia Commonwealth University, USA. In the validation process, the authors reviewed the 14 items in the physician version and eight of them were modified. The modified version was reviewed by a group 30 students and 30 members of academic staff for face and content validity. It was subsequently administered

to students in the first, second and third years of the course. A total of 652 students responded (the response rate not reported). After eight weeks, 10 students from each year were requested to complete the instrument again to compute the test-retest reliability. An exploratory factor analysis generated the same factors and items clusters as in the non-university physician version. The internal consistency and the test-retest reliability were acceptable (Cronbach alpha > 0.77). The scale scores helped compare clinical and non-clinical students; clinical students reported significantly better lifelong learning orientation compared to pre-clinical students (Wetzel *et al.* 2010).

d. Jefferson Scale of Empathy (Hojat *et al.* 2003a; Hojat *et al.* 2002; Hojat *et al.* 2001)

The primary focus of this scale is the measurement of empathy, which is an important aspect of professionalism of individuals and professionalism culture with regards to doctor-patient relationship and communication (GMC 2006; Jotkowitz *et al.* 2004; Swick 2000). It was also developed by a team led by Hojat at the Jefferson Medical School in the USA. It exists in three separate versions for students, trainees, and physicians. Unlike the lifelong learning scale, this was developed initially for students and residents, and subsequently validated to be used among physicians. In the development of the students and trainees version, the authors followed a method similar to that they followed for the lifelong learning scale described above. An extensive literature review was performed to identify the attributes of empathy (90 items), and a modified Delphi process with the participation of 100 physicians, deployed to condense the number of items to 45. In the first validation study, this scale was administered face-to-face in pen-and-paper format with a seven-point Likert type rating scale (1='strongly disagree' and 7='strongly agree') to

41 internal medicine residents and 193 third year medical students at the Jefferson Healthcare System. Based on a principal component analysis the number of items was reduced to 20 and three domains of empathy were identified; perspective taking, compassionate care, and ability to stand in the patient's shoes. The selected items demonstrated high internal consistency (Cronback alpha 0.87 for residents and 0.89 for students). The subsequent application of the 20-item scale and the analysis of results strongly supported the criterion-related and construct validity of the measure. A set of examples used in the students and resident version is included in Table 10. A validation study confirmed the relevance of the student and resident version in the UK context (Tavakol *et al.* 2011b). It was used in UK medical schools to investigate attitudes towards empathy and their differences by gender and stage of training (Austin *et al.* 2007; Tavakol *et al.* 2011a), and to compare emotional intelligence, academic performance, and empathy (Austin *et al.* 2007).

Table 10 - Domains and example items in the Jefferson Scale of Empathy (student and resident version) (Hojat *et al.* 2003a; Hojat *et al.* 2002; Hojat *et al.* 2001)

Domain	Example items
Perspective taking	A physician who is able to view things from another person's perspective can render better care
	For more effective treatment, physicians must be attentive to their patients' personal experiences.
Compassionate care	It is important to ask patients about what is happening in their lives as it is to ask about their physical symptoms
	Reading non-medical literature and enjoying the arts can enhance physicians' ability to render better care.
Ability to stand in the patient's shoes	Because people are different, it is almost impossible for physicians to see things from their patients' perspective.
	Emotion has no place in the treatment of medical illness

The physician version was adapted from the 'student & resident' version. Although the same items were used, the wording was changed to reflect empathic behaviours rather than attitudes. This was responded to by 704 physicians in the Jefferson Health System (response rate 70%) in a mail survey. The analysis of the results confirmed the high internal consistency of the scale (Cronbach alpha 0.81), the existence of the same three domains, and the appropriateness of the 20 items in the context of practising physicians.

Subsequent administration of this scale among students, residents and physician over the years generated a new insight into empathy in medicine: empathy among students and residents does not correlate with academic performance as measured by objective examinations, but it correlates well with global clinical competence judged by clinicians; in general, women always have more empathic attitude / behaviour than men; practitioners of 'people-oriented' specialties (e.g. family medicine, internal medicine, obstetrics and gynaecology) are more empathic than those who practice 'technology-oriented' specialties (e.g. surgery, anesthaesiology , radiology); and the empathic attitude declines progressively along the medical training (Hojat *et al.* 2003a).

e. Vignette-based Professional Attitude Questionnaire (O'Sullivan & Toohey 2008)

This questionnaire consists of 24 case vignettes of professionalism dilemmas to represent eight literature-based attributes of professionalism: (i) subordinate own interests to interests of patients, (ii) demonstrate high ethical and moral standards, (iii) behave according to an accepted social contact, (iv) demonstrate humanistic values such as integrity and honesty, (v) show responsibility and accountability, (vi) have a commitment

to improve, (vii) cope with complexity and uncertainty, and (viii) demonstrate reflective practice. It was responded to by faculty and undergraduate medical students in Australia. The vignettes for the faculty were identified by analysing and modifying incidents reported to a quality improvement committee of a hospital, and the vignettes for students were identified by analysing and modifying incidents reported to a university clinical teaching unit. These vignettes, therefore, appear to have been used as proxy measures of the same attributes. Examples case vignettes are included in Table 11. The vignettes were validated by 30 members of the academic staff for their appropriateness to represent each attribute. In each vignette, there were three behavioural options, which to be annotated with 'least likely to perform', 'most likely to perform' and 'not selected'. This pen-and-paper based questionnaire was responded to by 21 faculty members (response rate not reported) and 133 students from the second, fourth and sixth years of the course (response rate between 19 – 26%).

Table 11 – Examples of the professional attributes and relevant scenarios used in the Vignettes-based Professional Attitude Questionnaire (O'Sullivan & Toohey 2008)

Attributes	Relevant scenarios
demonstrate high ethical and moral standards	Scenario 10 You are a qualified medical officer who has been practising for 2 years. A Pharmaceutical Company Representative offers you a gift of a new stethoscope to remind you to prescribe their brand of medication instead of the generic brand (both medications are pharmacologically identical). You would: Option A: Accept the gift and agree to prescribe their medication when you remember to do so Option B: Decline the gift and point out to the Representative that gifts given with an intent to change prescribing behaviour are not appropriate Option C: Accept the gift and not agree to change your prescribing habits
demonstrate humanistic values such as integrity and honesty;	Scenario 12 You are an undergraduate medical student who is in the 6th year of a medical course. Prior to your final year Clinical Examination in Medicine you inadvertently come across the list of patients who have been invited to participate in this examination. Their diagnoses are also listed. You would: Option A: Return the list to the Clinical Teaching Unit (Education Unit) and alert them to how you obtained it inadvertently Option B: Destroy the list after you had read it and not discuss it with your fellow students Option C: Keep the list and distribute it to your study group

The authors used the faculty response as the benchmark. The vignettes helped identify the differences between and within student groups, and between the students and faculty. The second year students rated poorly in ethical and moral standards, humanistic values, and responsibility and accountability compared to faculty. The fourth and sixth year students, on the other hand, rated highly in ethical and moral standards, but lower than faculty in humanistic values. The authors identified these differences as gaps in professionalism education in their medical school. The concordance of responses between the faculty and students in all years was reported as high. Again, the overall scores of respondents can be used as an indirect measure of professionalism culture of the institution or programme.

f. Learning Environment for Professionalism Survey (Thrush *et al.* 2011)

Thrush *et al* (2011) developed this survey to measure the professional environment within clinical rotations of the undergraduate medical education programme, University of Arkansas, USA. In the development of the instrument, the authors adapted items from three previously-developed measures. One of the three studies selected focused on the measurement of the professional environment in the undergraduate context (Arnold *et al.* 1998) while the two remaining studies focused on the assessment of physicians in practice (Beaudoin *et al.* 1998; Szauter *et al.* 2003). All three studies were conceptually related to the ABIM framework of professionalism. The students were expected to indicate the frequency in which they observed the six professional and four unprofessional behaviours included in the instrument (Table 12), on a four-point scale (4 = 'consistently', 1 = 'never').

Table 12 - Examples of items included in the Learning Environment for Professionalism Survey (Thrush *et al.* 2011)

I have observed residents/attendings
• Who are positive role models of effective doctor–patient relationships.
• Inappropriately withholding information or intentionally giving incorrect information to a patient.
• Who were concerned about the overall well-being of patients, not just their presenting complaints.
• Treating non-physician healthcare workers in a disrespectful or inappropriate manner.
• Place the needs of their patients ahead of their own self-interest.

A total of 902 students responded to the survey over two years. A principal component analysis of the responses confirmed that the two groups of items corresponded to 'professionalism' and 'unprofessionalism' as intended. The details included in the paper

do not describe the exact procedure used for further analysis of the results and interpretation of findings. The authors, however, claimed that it was a valid and reliable tool to measure the professional environment in their clinical rotations.

g. Dundee Ready Education Environment Measure (DREEM) (Roff *et al.* 1997)

This instrument was developed to evaluate the educational environment of institutions training medical undergraduates. Although it is not a measure of professionalism culture, it deals with a similar construct and it has been used widely in different settings. Therefore, it is worth examining the DREEM in terms of the process of development and implementation.

The authors used a 'grounded theory' approach as there was no concrete definition of the 'learning environment'. Accordingly, student diaries, related issues raised in curriculum planning committees of the Dundee Medical School, UK, and a review of literature were used to develop the initial 110 items. Subsequently, these items were reduced to 50 with a two-iteration modified Delphi process. In numerous field studies conducted as pen-and-paper administrations, medical undergraduates around the world rated each item on a five-point Likert scale (strongly agree = four marks, to strongly disagree = zero marks). An exploratory factor analysis of responses generated five domains: students' perceptions of teaching (12 items), students' perceptions of teachers (11 items), students' academic self-perceptions (12 items), students' perceptions of atmosphere, and students' self-perceptions (seven items) (Table 13). The total and domain scores were used to comment on the nature of the educational environments.

Table 13 - Domains and examples of items of Dundee Ready Education Environment Measure (Roff *et al.* 1997)

Domain	Examples of items
Students' perception of teaching	I am encouraged to participate in class
	Long-term learning is emphasized over short term
Students' perceptions of teachers	The teachers are knowledgeable
	The teachers provide constructive criticism here
Students' academic self-perceptions	Much of what I have to learn seems relevant to a career in medicine
	I have learned a lot about empathy in my profession
Students' perceptions of atmosphere	There are opportunities for me to develop interpersonal skills
	The atmosphere motivates me as a learner
Students' social self-perceptions	There is a good support system for students who get stressed
	My accommodation is pleasant

This scale was validated and used not only in the UK but globally for different purposes, e.g. to make both cross-sectional and longitudinal institutional profiles of the learning environment, to compare institutions and the perceptions of different groups within the same institution, and to correlate students' academic performance and educational environment (Roff 2005). It has also been used among undergraduates outside medicine, e.g. dental, nursing and chiropractic (Roff 2005). DREEM demonstrated high internal reliability at both scale and domain levels across different settings (Riquelme *et al.* 2009). However, there are reports challenging its construct validity; all 50 items measure a single construct and the five domains are non-existent (Hammond *et al.* 2012).

h. Dundee Poly-professionalism Inventory I (Roff *et al.* 2011)

This online instrument focuses on the evaluation of the institutional culture of medical, dental and nursing schools in relation to the academic integrity. The development of this

measure is underpinned by the notion that fostering professionalism is a multi-stage process and the academic integrity is the first stage (Roff *et al.* 2009)

In the development of this instrument, 100 examples of academic dishonesty were identified by reviewing 30 cognate studies conducted around the topic. The authors subsequently condensed these to make 42 behavioural items (Table 14). The respondents, both students and members of the staff, indicated: (a) their involvement / engagement in such activities (on dichotomous yes / no scale), (b) their knowledge about such incidents in the institutional environment (on a frequency scale), and (c) sanctions for a student, who demonstrated such behaviours with no mitigating circumstances. A hierarchy of sanctions, adapted from Teplitsky (2002) and allocated one (least severe) to 10 (most severe) marks, was used as an indirect measure of the respondents' attitudes towards each behaviour (Table 15). The instrument was delivered online with the facility to self-generate basic descriptive statistics.

Table 14 - Examples of items included in the Dundee Poly-Professionalism Inventory I (Roff *et al.* 2011)

Plagiarising work from a fellow student or purchasing work from a supplier
Altering or manipulating data (e.g. adjusting data to obtain a significant result)
Inventing extraneous circumstances to delay sitting an exam
Engaging in substance misuse (e.g. drugs)
Lack of punctuality for classes
Examining patients without knowledge or consent of supervising clinician

Table 15 - Sanctions and corresponding scores used in Dundee the Poly-Professionalism Inventory I (modified from Teplitsky 2002)

Sanction	Corresponding score
None	1
Reprimand (verbal warning)	2
Reprimand (written warning)	3
Reprimand, plus mandatory counselling	4
Reprimand, counselling, extra work assignment	5
Failure of specific class/remedial work to gain credit	6
Failure of specific year (repetition allowed)	7
Expulsion from college (readmission after one year possible)	8
Expulsion from college (no chance for readmission)	9
Report to professional regulatory body	10

The pilot study of the Dundee Poly-Professionalism Inventory I was responded to by the faculty and students of the Dundee University Medical (28 staff, 375 students), Dental (16 staff, 79 students) and Nursing (13 staff, 218 students) Schools. The proposed sanctions demonstrated a significant attitudinal difference between students and the staff on academic dishonesty (Roff *et al.* 2011). Although not reported in Roff *et al*, 2011, in the analysis, it was realised that the first two questions about the involvement / engagement in and the incidence of behaviours did not provide useful and credible data, which could be interpreted meaningfully. It also revealed that the instrument has very high reliability (Cronbach alpha > 0.91), calculated using the sanction scores (Chandratilake *et al.* 2010). The final number of items was reduced to 30 after analysing the responses qualitatively (Chandratilake *et al.* 2010).

The key features of all the measures described above are summarised in the Table 16 below.

Table 16 - Comparison of scales which tend to measure quantitatively the professionalism culture / environment of medical schools

Instrument	The primary focus	No. and nature of items in the final instrument	Settings and respondents	Format	Type of scale used	Steps taken to establish validity	Determination of reliability	Practicability and acceptability
ABIM Scale (Arnold *et al.* 1998)	Assessing the professional environment using the ABIM framework as the basis	12, both professional and unprofessional behaviours	USA Both undergraduate and postgraduates Multicentre study	Pen-&-paper Postal	Nine-point scale unipolar	Uses ABIM as the basis An exploratory factor analysis of responses	Internal consistency	Not reported
PSCOM Professionalism Questionnaire (Blackall *et al.* 2007)	Assessing the professional environment using the ABIM framework as the basis	36, professional behaviours	USA Students and the faculty Single centre study	Pen-&-paper Self-administered	Five-point unipolar scale	Uses ABIM as the basis Delphi process Principal component analysis	Internal consistency	Not reported
Professional Climate Instrument (Quaintance *et al.* 2008)	Assessing the professional environment using the ABIM framework of professionalism as a basis	12, both professional and unprofessional behaviours	USA Fourth year undergraduates and the faculty Single-centre study	Pen-&-paper Self-administered and email	Four-point unipolar scale	Uses ABIM framework as the basis Focus group discussions with students to develop items	Internal consistency	Not reported
Medical Professionalism Attitudes and Beliefs Questionnaire (Kalet & Steven 2004)	Professional attitudes and beliefs of students	44, professional behaviours	USA Undergraduates Single-centre study	Pen-&-paper Self-administered	Details not available	Items developed from students' written accounts on professionalism Exploratory factor analysis	Internal consistency	Not reported

Instrument	The primary focus	No. and nature of items in the final instrument	Settings and respondents	Format	Type of scale used	Steps taken to establish validity	Determination of reliability	Practicability and acceptability
Moral Distress Survey (Wiggleton *et al.* 2010)	Professional environment via students' moral distress	55 , unprofessional events	USA Undergraduates Single-centre study	Online Email delivery	Five-point unipolar scale to indicate frequency of events and three-point bipolar scale to indicate the level of distress	Adapted a validated instrument in nursing supplemented by a literature review	Internal consistency	Not reported
Jefferson Scale of Lifelong Learning (Hojat *et al.* 2003b)	Lifelong learning orientation in physicians and students in an institutional environment	14, professional behaviours related to lifelong learning	USA Physicians and undergraduates Single-centre study	Pen-&-paper Postal	Four-point bipolar scale	A literature review Delphi process Exploratory factor analysis	Internal consistency Test-retest validity ? Construct validity	Not reported
Jefferson Scale of Empathy (Hojat *et al.* 2001)	Attitudes and orientation towards empathy among individuals in an institution / clinical environment	20, both empathic and non-empathic behaviours	USA Undergraduates, postgraduates and practitioners Single-centre study	Pen-&-paper Postal	Seven-point bipolar scale	A literature review Delphi process Principal component analysis Subsequent international validation studies	Internal consistency Criterion-related validity Construct validity	Not reported
Vignettes-based Professional Attitude Questionnaire (O'Sullivan & Toohey 2008)	Professional attitudes in general	24 professionalism dilemmas	Australia Undergraduates and the faculty	Pen-&-paper Face-to-face	Three options	A literature review, consensus of a panel	None	Not reported

Instrument	The primary focus	No. and nature of items in the final instrument	Settings and respondents	Format	Type of scale used	Steps taken to establish validity	Determination of reliability	Practicability and acceptability data
Learning Environment for Professionalism Survey (Thrush *et al.* 2011)	Professionalism environment in relation to ABIM domains	11, both professional and unprofessional behaviours	USA Undergraduates Single-centre study	Pen-&-paper Self-administered	Four-point unipolar scale	Used ABIM as the basis Items taken from previously validated instruments Principal component analysis	Internal consistency	Acceptability of the administration and the faculty is reported.
Dundee Ready Education Environment Measure (DREEM) (Roff *et al.* 1997)	Educational environment	50, positive and negative educational practices	UK Undergraduates	Pen-&-paper	Five-point bipolar scale	A literature review Students and staff feedback on curriculum Delphi process Subsequent international validation	Internal consistency	Not reported
Dundee Poly-professionalism Inventory I (Chandratilake *et al.* 2010; Roff *et al.* 2011)	Environment related to academic integrity	30, items related to academic dishonesty	UK Undergraduate Faculty	Online Delivered by email	Ten-point sanctions scale	A literature review	Internal consistency	Not reported

2.4.3. Discussion

The insight into the measurement of the professionalism culture gained by the literature search and the analysis of existing measures is discussed below.

2.4.3.1. Geography of origin

Most of the measures of professionalism culture / environment found in the literature originate from the USA and the UK. The renewed interest in and emphasis on professionalism is still gaining momentum in other parts of the world (Hodges *et al.* 2011). Although there were several qualitative studies of institutional environment in relation to different aspects of professionalism in both academic (e.g. Lempp & Seale 2004) and clinical (e.g. Monrouxe *et al.* 2011) settings from the UK, the quantitative measurements were found to be less common than from the USA. This may well be attributed to the fact that, in the USA, there is a well-demarcated and nationally-accepted 'definition' for professionalism (ABIM 2000), but, in the UK, there is no such definition except for a set of expectations proposed by the GMC (GMC 2006). The existence of a clear definition or theoretical conceptualisation for the construct concerned facilitates the process of developing survey instruments in general (DeVellis 2003, pp. 1-13) and for professionalism in particular (Baldwin & Daugherty 2006). The preference for a qualitative approach to the topic among UK medical educationalists does not necessarily reflect a general scepticism towards quantitative approaches. In assessing the professionalism of individuals, the Royal Colleges in the UK recommend and use quantitative assessments such as multi-source feedback, developed originally in the USA and subsequently adapted to the UK context (RCP 2009).

2.4.3.2. The focus

It is evident that certain measures, e.g. ABIM scale, Professionalism Climate Instrument, Moral Distress Survey, Vignettes-based Professionalism Attitudes Questionnaire, were designed to capture the concept of professionalism culture / environment as a whole while some others were focused on a particular aspect of professionalism, e.g. Jefferson Scales of Lifelong Learning and Empathy, Dundee Poly-professionalism Inventory 1 (Table 16). The presence of a clear definition for professionalism may be one of the helpful factors for the development of measures to capture professionalism as a whole. However, in developing a measure, whether to consider professionalism as a whole or from its individual components depends very much on the research question (DeVellis 2003). Given the nature of its content, the focus of the DREEM was on measuring the educational environment rather than the professionalism environment of an institution. However, apart from its content, the DREEM did provide useful insights into the development of a measure for the institutional environment.

Certain measures had defined or redefined the concept of professionalism or its elements. For example, with the Jefferson Scales of Lifelong Learning and Empathy, it was measurement of these constructs in the institutional environments which generated widely-accepted working definitions for the relevant constructs. The Professional Climate Instrument, on the other hand, redefined professionalism in the USA context by identifying new dimensions.

2.4.3.3. The items

While various instruments have used positive and negative statements only and a mixture in same, none of the authors have reported the rationale behind their preference of one statement type over the other. However, it is the general belief that mixing positive and negative statements reduces the agreement bias, i.e. bias caused by respondents randomly selecting responses without considering the content (DeVellis 2003, pp. 69 & 70; Schuman & Presser 1996, pp. 203 - 230). However, contradicting this popular assumption, the empirical research in psychology and sociology has demonstrated that this bias has almost no bearing on positive or negative orientation of statements (DeVellis 2003, pp. 69 & 70), but on other factors such as the ignorance of respondents on the topic and ambiguity of the statements (Schuman & Presser 1996, pp. 203 - 230). In addition, DeVellis (2003, p.69) argues that reversal of item polarity in the same survey instrument may make respondents confused about 'the difference between expressing their strength of agreement with a statement, regardless of its polarity, and expressing the strength of attribute being measured' (e.g. professionalism). In addition, statistically, the negatively and positively worded items may simply cluster into two factors, regardless of the relationship between these items (Meade & Craig 2012). The work of Hodges *et al* (2009) and Ainsworth *et al* (2006) suggest that students, trainees or medical professionals are better at identifying unprofessional than professional behaviours. In contrast, Dewitt and Daugherty (2006, pp.95 - 112) have observed that, in the assessment of professionalism (at individual level as well as institutional level), using unprofessional behaviours, (popular in the 1980s) has been replaced by the use of professional behaviours with the attainment

of explicit insight into what is meant by professionalism. The current trend, therefore, is the encouragement towards and the rewarding of professional behaviours via educational interventions such as the measures discussed above, leaving the unprofessional behaviours to be identified through reporting systems and dealt with by disciplinary committees (Baldwin & Daugherty 2006, pp. 95 - 112).

For any measurement of professionalism to be effective, professionalism and its components should be operationalised in behavioural terms (ABIM 2000). This phenomenon is well-reflected in all the measures under discussion. However, the orientation of these behaviours in different instruments varies as the wording of items and rating scales used are different. In the Jefferson Scale of Lifelong Learning and Vignettes-based Professional Attitude Questionnaire the respondents are expected to report the extent to which they behave as described in a given item. In the ABIM Scale, Professionalism Climate Instrument and Learning Environment for Professionalism Survey, the respondents are expected to report the frequency of bahaviours observed in others. The Moral Distress Survey contains items, which are focused on the respondents' own behaviours and the behaviours observed in others; the respondents are expected to report how common they are. The behavioural items included in the PSCOM Professionalism Questionnaire, Jefferson Scale of Empathy, Dundee Poly-Professionalism Inventory I and the DREEM expect the respondents to report neither their own behaivours nor the behaviours observed in others, but their attitude towards such behaviours. As the study reports no details, the orientation of items in Medical Professionalism Attitudes and Beliefs Questionnaire cannot be commented on. Baldwin and Daugherty (2006, p.107)

claim that, under proper conditions, students and residents accurately report even their own unprofessional acts in surveys. However, there is more evidence to suggest that the responses are more credible and reliable when the items are oriented towards behaviours of peers or other members of the team than towards respondents' own behaviours (Eva & Regehr 2005; Rees & Shepherd 2005b; Wilkinson *et al.* 2009). Even reporting the behaviours of others may be influenced by the level of morality of respondents themselves; a student with high morality may observe a higher number of unprofessional behaviours among their colleagues than others (Branch 2000). The suitability of attitudes and behaviours in measuring an institutional environment or culture is discussed in detailed in a subsequent section (Section 2.5, p.1).

The number of items included in various measures vary from 11 (Learning Environment Survey) to 55 (Moral Distress Survey). The greater the number of items in the instrument, the higher the internal consistency and sampling validity (DeVellis 2003 pp. 27-38). The literature on the relationship between the length of a survey instrument, i.e. the number of items, and the response rate is inconclusive (Sheehan 2001). However, fewer items put less burden on respondents, which may enhance the credibility of responses and, therefore, a balance between brevity and reliability is recommended (DeVellis 2003, pp.96-100). Although there are instances where high internal consistency values were achieved with a lesser number of items, at least around 20 survey items are necessary to adequately sample a given construct, i.e. to achieve content validity (Clark & Watson 1995; Kline 1993, pp. 530 - 550). If a high internal consistency with a few items (e.g. 10) has been achieved, the items in the measure are most likely simple paraphrases of each

other, argues Kline (1993, p. 532). In that sense, the ABIM Scale, Professional Climate Instrument, Jefferson Scale of Lifelong Learning and Learning Environment for Professionalism may possess an inadequate number of items to meet the validity requirements. Factor analysis of responses, however, provides supportive validity evidence (Clark & Watson 1995; Hojat *et al.* 2003a). As the items in the ABIM Scale and the Jefferson Scale of Lifelong Learning cluster under a meaningful factor structure, the threat to content validity posed by a reduced number of items appears to be low. The factor analysis conducted for the Learning Environment for Professionalism, however, did not generate a meaningful factor structure, which questions its validity to measure the intended construct.

Different authors have used different methods in reducing the number of items; statistical methods such as factor analysis, e.g. ABIM Scale, Jefferson Scales of Empathy and Lifelong Learning, or qualitative methods, e.g. Dundee PolyProfessionalism Inventory I. The advocates of survey development (Clark & Watson 1995; DeVellis 2003; Kline 1993), however, advise the use of statistical methods such as factor analysis over qualitative methods. They advise the use of the former as a guidance rather than a rule. For a factor to be statistically stable (reliable), it should be represented by at least two items (DeCoster 1998). The measures under discussion comply with this criterion.

2.4.3.4. Validity and reliability evidence

In the development of measures, three initial steps are recommended to establish the content validity: (i) the definition / conceptualisation of the desired construct, (ii) the

identification of as many attributes (or items) related to the construct as possible, and (iii) the selection of items for the measure with the consensus of experts (DeVellis 2003, pp.60 - 63; Johnston & Wilkinson 2009). These three steps also help establish the substantive component of construct validity of the instruments. After completing these steps, the items should be piloted with an adequate number of respondents (DeVellis 2003, p.88). The statistical evaluation of items (e.g. internal consistency and factor analysis) based on the responses after this field-test provides supportive evidence for the content validity and for the structural component of the construct validity (Clark & Watson 1995). These steps have been followed in the development of the PSCOM Professionalism Questionnaire, the Jefferson Sales of Empathy and Lifelong Learning and the DREEM. However, they have not been followed with the same strength in the development of the ABIM Scale, Professional Climate Instrument and the Medical Professionalism Attitudes and Beliefs Questionnaire, Moral Distress Survey or the Vignettes-based Professional Attitude Survey. Out of all the measures, the Learning Environment for Professionalism Survey and the Dundee Poly-Professionalism Inventory I followed these steps poorly.

Most of the instruments discussed above have reported the internal consistency at both scale and domain levels. The values are above the acceptable levels which indicate that the items represent a single construct of one of its domains (Kline 1993, pp.27 - 39). However, these high values should be correctly interpreted; a large number of poorly related items or a small number of paraphrasing items can also increase the internal consistency (Kline 1993, pp.90 - 100). Although, if properly interpreted, high internal reliability is a positive indicator of the psychometric rigour of an instrument, the evidence

of external reliability (test-retest and parallel form reliabilities) should also be reported (Kline 1993, pp.5-14). Only the Jefferson Scales have provided literature evidence for test-retest reliability, which are above acceptable levels (Kline 1993, pp.5-14). However, at the development stage of an instrument, reporting internal consistency may suffice (Kline 1993, pp.5-14). It appears that, out of all instruments, the two Jefferson Scales have provided the most evidence for validity and reliability.

The main drawback of self-reported measures about professionalism in terms of validity is that they gather the perceptions of the respondents, which may not reflect their actual behaviours (Baldwin & Daugherty 2006, p.102). Ideally, the findings of these instruments should be cross-validated with more objective methods, such as observations (Baldwin & Daugherty 2006, p.102). The measurement of attitudes, as done in most of these instruments, may not accurately predict behaviours; attitudes are only one of the multiples factors leading to a behaviour (Ajzen 1991).

2.4.3.5. Respondents and response rate

The Medical Professionalism Attitudes and Beliefs Questionnaire, the Moral Distress Survey, the Learning Environment for Professionalism Survey and the DREEM have made conclusions about the professional environment or component of it based on the perceptions of undergraduates alone. The ABIM Scale has utilized the responses of both undergraduates and postgraduates. The PSCOM Professionalism Questionnaire, the Professionalism Climate Instrument, both the Jefferson scales, the Dundee Poly-professionalism Inventory I and the Vignettes-based Professional Attitude Questionnaire

have used both students and faculty. The collation of perceptions of all stakeholder groups provides more credible results than a single source (Baldwin & Daugherty 2006; Rees & Shepherd 2005a). On the other hand, students, postgraduate trainees, faculty and clinicians contribute in their own way to the institutional environment, which should be taken into account in understanding professionalism at institutional level (Hafferty 1998). Therefore, the instruments which gathered the perceptions of both students and teachers may measure the professional environment more credibly than the rest.

An adequate number of responses to the first field-test is essential to statistically examine the internal structure of a given instrument (e.g. factor structure, internal consistency) though there is no agreement on the number of survey items versus responses needed (DeCoster 1998; DeVellis 2003, p.88). However, the sample should be 'sufficiently large to eliminate the subject variance as a significant concern' (DeVellis 2003, p.88). It is recommended that an average scale (20 – 30 items) requires around 200 – 300 responses in the first field test (DeCoster 1998; DeVellis 2003, p.88). All instruments discussed above have reached well above this recommendation in their respective field-tests, indicating the reliability of statistical analysis.

The response rate, however, is important to allow correct interpretation of statistical comparisons between demographic sub-groups of respondents. The response rates of the first field tests were reported in all the above studies except the Professionalism Attitude and Beliefs Questionnaire, Learning Environment for Professionalism, and the DREEM, which were delivered face-to-face. The ABIM Scale, which was distributed face-to-face but

returned by mail, achieved an overall response rate of 75%. The physicians' version of the Jefferson Scale of Lifelong Learning was completed by 58% of faculty and 43% of practitioners. Its student version, administered face-to-face, achieved a response rate of 89%. However, the PSCOM Professionalism Questionnaire (response rate 51%) and Vignettes-based Professional Attitude Questionnaire (response rate 19 – 26%), which were entirely self-administered, achieved comparatively lower response rates than similar measures. In the US context, the average response rate for a postal questionnaire among medical professionals is 51% (Cook *et al.* 2009). Therefore, the response rates of the instruments delivered via post in the USA are reasonable. The Professionalism Climate Questionnaire achieved a response rate of 68% by administering it face-to-face and by sending an email version. This approach to administration, which is referred to as mixed model design, attempts to boost the response rate (Mitra *et al.* 2008). The Moral Distress Survey (response rate =60%) and the Dundee Poly-professionalism Inventory I (response rate between 22 – 50%) were delivered online. Surveys administered online among US undergraduates tend to secure 10% fewer responses than mail surveys (Mitra *et al.* 2008). Only around 30% of Canadian medical professionals (including medical students) usually respond to online surveys (Grava-Gubins & Scott 2008). Therefore, a response rate above 40%, which has been reached by both online instruments under discussion, appears to be an achievement.

2.4.3.6. Rating scales used

The instruments under discussion, except for the Dundee Poly-professionalism Inventory I and the Vignettes-based Professional Attitude Questionnaire, use Likert-type rating scales, but with a different number of response categories and bipolar or unipolar variation. The number of response categories varies from nine (e.g. ABIM Scale) to three (e.g. Vignettes-based Professional Attitude Questionnaire). In addition to the number of items in an instrument, the number of response categories in the rating scale determines the validity and reliability of the instrument (DeVellis 2003, pp.70-85). Usually there is an inverse relationship between the number of items and response categories; the higher the number of items, the fewer the number of response categories required to obtain discriminative responses (DeVellis 2003, pp.70-85). However, fewer than four response categories demonstrate no significant discrimination while more than seven categories add no significant improvement in discrimination between items (Lozano *et al.* 2008). Therefore, for example the ABIM scale may have produced the same or similar results with five response categories. However, the number of response categories used by all measures under discussion has helped discriminate responses except for the Vignettes-based Professional Attitude Questionnaire. Given the number of items (24) and the number of response categories (three) the discrimination of responses may not be credible. Unlike other measures, in the Poly-professionalism Inventory I, a sanction code has been used as a measure of the attitude strength. Of course, the assumption made in the measure that the sanctions form a hierarchy, may not necessarily be accurate. Therefore, they may not essentially indicate the strength of the respondents' attitudes;

the implicit measurement of attitude, such as the sanction code, may carry problems of interpretation (Ranganath *et al.* 2008).

A unipolar rating scale can be used with an odd or even number of response categories as it needs no middle position (DeVellis 2003, pp.70-85) . However, a bipolar rating scale with an even number of response categories may force choice; the respondent must commit to either the positive or negative side of the scale as there is no middle position (DeVellis 2003, pp.70-85). On this basis, the use of a four-point bipolar scale may have caused bias to a certain degree in the Jefferson Scale of Empathy.

2.4.3.7. Practicability and acceptability evidence

The practicability, i.e. the ease of use, and acceptability, are two important determinants of the utility of a measurement instrument (van der Vleuten 1996). No instrument, except for the Learning Environment for Professionalism has reported its acceptability. Repeated use of the instrument in multiples centers and countries, i.e. the generalisability, however, may provide indirect evidence for its utility (Roff 2005). The Jefferson Scales of Empathy and Lifelong Learning, and the DREEM have been translated into many languages and used in different countries, which may indicate their acceptability and practicability. However, if an instrument is introduced to measure the professionalism culture as a part of the educational programme, at least the reaction of respondents should be evaluated to determine its acceptability (Hutchinson 1999; Rees & Shepherd 2005a).

2.4.4. Summary

The historical instruments for measuring the professionalism culture of medical schools quantitatively have their origin mainly in the USA. These instruments focus on professionalism as a single concept or components of it, and are based on either established definitions of professionalism or self-defined frameworks of professionalism. They measure either attitudes or behaviours though both should be measured in order to make credible conclusions on the institutional professionalism culture. The number of items included in every instrument under discussion was adequate to represent the construct and achieve acceptable reliability. The steps followed to establish the validity however, varied between instruments. Although the numbers of responses achieved by the instruments, which have established their structural validity through statistical means, were satisfactory, the response rates achieved by some were insufficient to comment on demographic differences. Most instruments use Likert-type rating scales with an acceptable number of response categories. Despite the importance, the practicability and the acceptability of most of the instruments have not been reported.

This section of the literature review draws the following conclusions.

- Having a concrete definition of professionalism in the given context (e.g. ABIM definition in the US) is advantageous in developing quantitative measures of professionalism culture. However, as shown by many researchers (Hojat *et al.* 2001; Hojat *et al.* 2003b; Roff *et al.* 2012; Roff & Dherwani 2011) measures of professionalism and related areas can be developed even in the absence of such definitions.

- In the absence of a well-accepted definition, a validated quantitative measure of the professionalism culture of a medical school might be useful and acceptable to the UK context.
- Developing a large pool of items to represent the desired construct, logically reducing them to a manageable number via a consensus process, and deciding the final number of items via statistical methods are the recommended and commonly followed steps in establishing content validity.
- Although the use of professional behaviours as items constitute the current 'trend', the use of unprofessional or a combination of professional and unprofessional behaviours may make no difference in statistical analyses of results. The number of items included in the final measure, however, should be adequate to represent the desired construct, but should produce reliable results and be acceptable to respondents.
- The rating scales yield maximum effective results with four to seven response categories. Statistical interpretations of responses, such as reliability and factor analysis provide useful guidance on validity and reliability of the measures, and in managing the number of items in the final product.
- The measurement of respondents' perceptions may not necessarily reflect their actual behaviours. This may be improved by obtaining the perceptions of different groups interacting in the same environment.
- The practicability of the measures to be used in a given context and their acceptability to the respondents are important if such measures are planned to be used as a part of an educational programme.

None of the instruments discussed above has used a theoretical basis to minimise the possible mismatch between perceptions and reality and none of them has addressed the issue of establishing the relationship between attitudes, behaviours and social pressures, which are the constituents of any professionalism culture (Baldwin & Daugherty 2006; Rees & Knight 2007). Therefore, identifying an appropriate theoretical framework to address these issues was considered a meaningful and necessary next step.

2.5. Establishing the theoretical relationship between attitudes, social pressure and behaviour

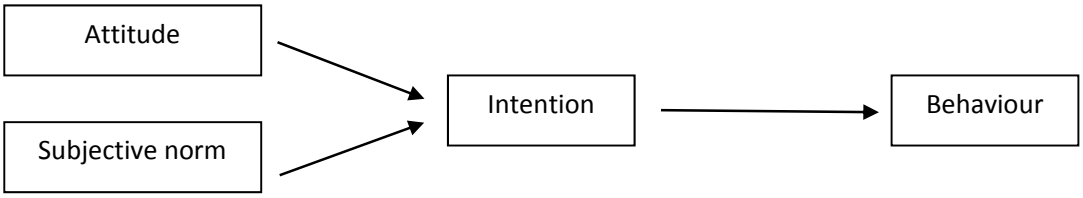
There are several psychological theories, which attempt to establish the relationship between attitudes, social pressure and human behaviour. However, theories hypothesising general dispositions, such attitude and personality trait, as predictors of human behaviour in specific situations have constantly demonstrated a poor predictive validity (Ajzen 1991; Rees & Knight 2007). The Theory of Reasoned Action (TRA) and its extension as the Theory of Planned Behaviour (TPB) have been shown to be useful in predicting health-related behaviours (Sheeran *et al.* 2003).

2.5.1. Theory of Reasoned Action (TRA)

According to the TRA introduced by Fishbein and Ajzen in 1975, the behaviours, over which individuals have control (i.e. voluntary behaviours), are performed by individuals only if they have the intention, i.e. people do things that they intended and do not do what they did not intend (Sheeran *et al.* 2003; Sheppard *et al.* 1988). The intention, which is the cognitive representation of one's readiness to perform a behaviour, is determined by one's attitude and subjective norms on the behaviour concerned (Sheeran *et al.* 2003).

'Attitude is the person's overall evaluation of what it would be like to perform a particular behaviour whereas subjective norm is the person's perception of social pressure to perform the behaviour' (Sheeran *et al.* 2003).(Figure 4)

Figure 4 – The relationship between attitude, social norm, intention and behaviour described in the Theory of Reasoned Action



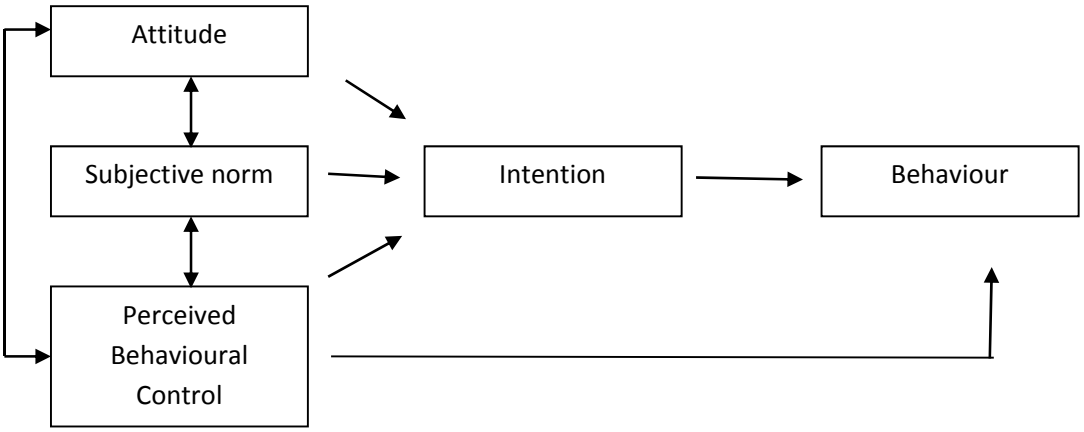
Sheppard *et al* (1988), in their meta-analysis on the research based on the TRA, conclude that there is strong overall evidence of the predictive utility of this theoretical model. In studies conducted prospectively, i.e. measurement of intentions followed by the measurement of behaviours; the intentions explain up to 38% variance of the respective behaviours; and attitudes and subjective norms explain up to 50% of variance in intentions (Sheeran *et al.* 2003). However, as defined in the theory itself, the applicability of TRA is limited to voluntary behaviours. As argued by Sheeran *et al* (2003), one might give up performing a strongly intended behaviour if one is short of the necessary ability, resources or opportunity. Therefore, when certain behaviours are concerned, the degree of control one has over the behaviour will be another determinant of its performance or non-performance (Sheeran *et al.* 2003). In such situations, the TRA will not reliably establish the intention-behaviour relationship. This has implications for investigating behaviours of healthcare workers as, in most instances, the majority of such workers do

not have total control over their professional behaviours in the healthcare environment (Randall & Gibson 1991; Walker *et al.* 2004).

2.5.2. Theory of Planned Behaviour (TPB)

The Theory of Planned Behaviour (TPB) was introduced to overcome the above limitation of the TRA (Ajzen 1991). The TPB hypothesised that, in addition to attitude and subjective norm, one's perception of one's control over the performance of the behaviour concerned, termed as Perceived Behavioural Control (PBC), is a variant of determining intention (Ajzen 1991). PBC was proposed also as an independent predicator of behaviour (Ajzen 1991) (Figure 5).

Figure 5 – The relationship between attitude, social norm, perceived behavioural control, intention and behaviour as described in the Theory of Planned Behaviour



The predictive validity of TPB depends on three conditions, namely: (1) the intention and the PBC should correspond to the same behaviour; (ii) intentions and PBC should remain stable between the assessment of intentions and behaviour; and (iii) PBC should be as accurate as actual behavioural control (Ajzen 1991). Inclusion of PBC may alleviate the

limitation of needing complete voluntariness of the behaviour for predicting its performance, but may not completely overcome this limitation as PBC will not correspond completely to the actual behavioural control (Ajzen 1991). However, in their meta-analysis on studies focused on predicting behaviours using TPB, Armitage and Conner (2001) conclude that, on average, the intentions- behaviour correlation is 40% while intention and PBC - behaviour multiple correlation is 52%. This meta-analysis confirms the validity of the theory by establishing statistical correlations between; attitudes, subjective norms and PBC versus intention, and PBC versus intention (Armitage & Conner 2001).

Since its inception, TPB has been used with reasonable accuracy in predicting and explaining a wide range of practices and behaviours of healthcare professionals. Bombeke *et al* (2010) used this theory to analyse patient-centred behaviours of medical students in Belgium; the theory provided a useful framework to ground their qualitative data. The TPB was proposed as a model of assessing and fostering professional behaviour among students and practitioners (Archer *et al.* 2008; Rees & Knight 2007). In addition, it was used in analysing and predicting behaviours and cultures of many health related behaviours. According to McMillan and Conner (2009), the prediction of the alcohol and tobacco use of 494 Leeds undergraduates by the TPB variables, i.e. intentions and PBC, was significant. Behaviour of ethical decision making of USA nurses was also predicted using TPB (Randall & Gibson 1991). The study reported that intentions and PBC together explain a significant component of variation in the desired behaviour. Walker *et al* (2004) reported that the intention and behaviour relationship described in the TPB was confirmed in their study on pharmacists' behaviour in treatment of vaginal candidiasis

with non-prescribed medicines. In the USA, differences in implementation intentions of 78 mental health practitioners, who learnt about new techniques in psychiatric management in a continuing medical education course, explained 12% variance of subsequent practice (Casper 2008).

It can be concluded that the TPB can be used to establish and explain the relationship between attitudes, social pressures and subsequent behaviour of healthcare workers, which is a helpful basis for measuring professionalism culture.

2.6. Summary of literature review

Professionalism is the characterisation of medicine as a profession. It evolves not only with the advancement of medicine as a science but also with socio-societal expectations. Therefore, professionalism is a dynamic and culture-sensitive concept, and has no universally agreed definition. However, most people agree that it encompasses attitudes, values and behaviours of doctors involved in the delivery of healthcare.

Professionalism has become a prime focus of medical education in recent years. It has been demonstrated to be a construct, which can be taught and learned. However, the nature of the construct itself is a challenge to medical educators, who have used variety of methods for both teaching / learning and assessment, under the umbrella of a formal curriculum. The learning goals of formal professionalism education appear to form a hierarchy from imparting the appropriate cognitive base to inculcating professional attitudes and behaviours. A large component of professionalism, however, is learned by students through the informal and hidden curricula. A major determinant of these

curricula is the professionalism culture of an institution. Therefore, the professionalism culture plays an influential role in the formation of professional identity of students as doctors. Exploring and understanding the institutional professionalism culture is an important strategy of professionalism education.

The professionalism culture of an institution is the cumulative manifestation of the attitudes, social pressures and behaviours of its inhabitants. It has shown to be measured quantitatively. However, in the UK, no quantitative measures of the professionalism culture in medical schools have been developed to date. There are several instruments developed for this purpose, especially in the USA. The psychometric quality of these measures, however, varys. They lack the theoretical basis to minimise the mismatch between the perceptions of the respondents and realism, and to establish the relationship between the different components of professionalism culture. The TPB appears to be an appropriate theoretical framework to develop ways of measuring the professionalism culture.

Therefore, to develop a quantitative measure of professionalism culture in the UK, the validity should be established by operationalising the concept of professionalism acceptable to the UK context and by basing the measure on theory of planed behaviours. The new measure needs to be tested subsequently for its reliability and acceptability.

2.7. Research questions and objectives of the study

Based on the literature review the following research questions were formulated:

1. How can ‘medical professionalism’ be made a measurable construct in the context of UK institutional culture?
2. How can the institutional professionalism of medical schools in the UK be measured quantitatively with appropriate validity, reliability and acceptability?

These research questions were addressed in the following objectives:

1. To conceptualise the term ‘medical professionalism’ and identify its attributes appropriate and acceptable to the UK context;
2. To develop a content-valid instrument to measure the professionalism culture of UK medical schools; and
3. To evaluate the reliability and the practicability of the instrument.

The project was conducted and presented in two phases. The two research questions were addressed sequentially in each of the phases.

Section B: Establishing the content validity of the measure of institutional professionalism culture

In this section, the research question 1, 'how can medical professionalism be made a measurable construct in the context of UK institutional culture?' is addressed. This research question focuses on the establishment of content validity of the measure of professionalism culture.

Content validity is defined as the degree to which elements of an assessment or a measurement instrument are relevant to, and representative of, a targeted topic (Hayens *et al.* 1995). Establishing content validity is usually a deductive process, i.e. defining virtually all items belonging to the targeted domain and sampling these systematically and representatively for the desired measurement instrument (Cronbach & Meehl 1955). The primary requirement of 'systematic deduction' for establishing content validity, therefore, is the selection of items with stakeholder consensus. This section is, therefore, organised into two chapters:

- **Chapter 3**: discusses how the attributes of professionalism in relation to the UK undergraduate medical context were identified and made measureable.
- **Chapter 4**: discusses how these attributes were validated using the consensus of the general public and medical professionals.

Chapter 3 – Identifying attributes of professionalism

3.1. Introduction and rationale

Several experts suggest (Clark & Watson 1995; DeVellis 2003 pp.63 - 65) that the identification of the attributes to represent the construct (professionalism in this instance), and making these attributes measurable should be the first stage in the development of measures. A number of researchers (Arnold *et al.* 1998; Hojat *et al.* 2003b) have followed this process in developing measures of professionalism culture. Such attributes have been identified by several methods: using the existing attribute framework as the basis (Arnold *et al.* 1998; Blackall *et al.* 2007; Quaintance *et al.* 2008); reviewing the literature (Hojat *et al.* 2001; Hojat *et al.* 2003b; Roff *et al.* 2011; Wiggleton *et al.* 2010); analysing expressions of key stakeholders, e.g. students (Kalet & Steven 2004); or combining these methods (Roff *et al.* 1997).

The ABIM framework of professionalism was formulated through an extensive literature review with the purpose of operationalising the abstract concept of professionalism, primarily for educational purposes (ABIM 2001). Therefore, in the scale development, the ABIM framework has provided guidance for narrowing down 'professionalism' to a measurable construct, and ensured its context-appropriateness (Quaintance *et al.* 2008). Subsequent analysis of responses to the scales has supported the construct validity of this framework to a great extent (Arnold *et al.* 1998; Quaintance *et al.* 2008). However such a framework may not necessarily encompass the concept of professionalism in its entirety (Arnold *et al.* 1998; Quaintance *et al.* 2008). The GMC's *Good Medical Practice*, on the

other hand, is primarily a set of guidance to encourage professional behaviour and discourage unprofessional behaviours among doctors (GMC 2006), and unlike the ABIM framework it has not attempted to operationalise 'professionalism' in the UK context. Therefore, using the GMC framework alone to conceptualise professionalism in the UK context may leave gaps (Brockbank *et al.* 2011; Horton 2005).

A review of the literature provides a sound basis to conceptualise and make the previously undemarcated concepts measurable (Clark & Watson 1995). In the absence of a concrete and well-accepted definition for the construct concerned, some authors (Hojat *et al.* 2001; Roff *et al.* 2011; Wiggleton *et al.* 2010), have used this approach effectively in developing psychological measures. However, repeated administrations of the scale and statistical analyses are necessary to establishing the ultimate validity of such scales for the measurement of the relevant constructs such as professionalism (Hojat *et al.* 2002).

The inputs of stakeholders, e.g. reflective accounts of students, have been used to define and operationalise professionalism (Kalet & Steven 2004). However, as the perceptions of different stakeholder groups, e.g. students versus physicians versus patients, may be different (Brockbank *et al.* 2011), triangulation of perceptions of different groups would be necessary for improved credibility.

The optimal approach to identifying and operationalising 'professionalism' would, therefore, appear to be the combination of the multiple strategies discussed above. In developing the DREEM, Roff *et al* (1997) triangulated several strategies; considered previous conceptual frameworks of the academic environment, reviewed the literature,

and analysed the feedback from both students and teachers. This may have helped the DREEM to become a generic measure, which successfully crossed different contextual and geographical boundaries.

3.2. Method

In the UK context, there is no well-accepted definition for professionalism (Brockbank *et al.* 2011; Horton 2005). Therefore, to conceptualise and make professionalism measurable the following steps were followed as informed by the literature.

1. The GMC *'Good Medical Practice'* was used as the basis.
2. The attributes identified by the GMC was supplemented with
    a. related and well-established frameworks of professionalism
    b. key studies conducted mainly in UK, Europe and North America.
3. A group of medical educators reviewed the final list of attributes for clarity, exclusivity (i.e. not multiple attributes conveying the same meaning) and relevance to undergraduate context.

In the review of literature to identify the attributes of professionalism as appropriate to the current context and the aim of the project, it was evident that the majority of studies have stemmed from either well-accepted professional / regulatory frameworks of professionalism (e.g. ABIM) or a few key definitions. Therefore, the articles which reported the underpinning frameworks for multiple studies and articles with special relevance to the project were selected to for use. The rationale for selecting each source is discussed below.

3.2.1. The rationale for selecting the sources

3.2.1.1. Guidelines of regulatory and professional bodies

'*Good Medical Practice*' published by the GMC (2006) was used as the starting point, as the GMC is the sole regulatory body of medical practice in the UK. The GMC professionalism guidelines are applicable to all medical practitioners, and they also are a framework of accountability for alleged unprofessional behaviour (GMC 2006). It also provided an insight into the regulator's perspective of 'professional' behaviour. The Royal College of Physicians (RCP), UK, on the other hand, is a leading professional body, which strives to maintain the professional standards of medical practice not only in the UK but also around the world through educating and supporting their membership throughout their careers (RCP 2005). Although there are several Royal Colleges for various medical disciplines, no other Royal College has published a dedicated report on professionalism. The report by the RCP, furthermore, has been well-publicised and widely-discussed in the literature (Wass 2006) and has been quoted as representative of the UK conceptualisation of professionalism (e.g. Cruess & Cruess 2006; van Mook *et al.* 2009c). It is valid and relevant to the modern context of professionalism since the mandate given to the RCP working party was redefining professionalism in a changing world (RCP 2005; Wass 2006).

The UK frameworks were supplemented with two international conceptual frameworks; the ABIM framework of professionalism (ABIM 2001) (Table 2), and the professionalism domain of the CanMEDS competency framework formulated by the Royal College of Physicians and Surgeons, Canada (RCPSC 2005) (Table 17). Both have emerged from North America specifically to operationalise professionalism in order to teach and assess this

abstract construct (Brennan *et al.* 2002; Frank & Danoff 2007). Historically, there may be differences in conceptualising professionalism between the British and North American medical communities (Hafferty 2006). However, with globalisation of healthcare and global appreciation of patient-centredness at the heart of professionalism in todays' context, it is worth considering these frameworks to widen the scope. Furthermore, these frameworks have been validated in many countries and found to be mostly relevant and valid. For example: the ABIM framework has been used in Taiwan (Tsai *et al.* 2007) and in Iran (Aramesh *et al.* 2009); the validity of the CanMEDS framework has been demonstrated in Denmark (Ringsted *et al.* 2006) and the Netherlands (Verkerk *et al.* 2007). As discussed in the literature review, several measures of institutional professionalism (Table 16, p.1) (Arnold *et al.* 1998; Blackall *et al.* 2007; Quaintance *et al.* 2008; Thrush *et al.* 2011) have successfully used the ABIM framework as the basis for item identification or development.

Table 17 - Elements of 'Professional' domain of CanMeds roles (RCPSC 2005)

Altruism
Integrity and honesty
Compassion and caring
Morality and codes of behaviour
Responsibility to society
Responsibility to the profession, including obligations of peer review
Responsibility to self, including personal care in order to serve others
Commitment to excellence in clinical practice and mastery of the discipline
Commitment to the promotion of the public good in health care
Accountability to professional regulatory authorities
Commitment to professional standards
Bioethical principles and theories
Medico-legal frameworks governing practice
Self-awareness
Sustainable practice and physician health
Self-assessment
Disclosure of error or adverse event

3.2.1.2. Key studies in professionalism

After identifying the key attributes of professionalism included in the frameworks formulated by the key local and international regulatory and professional bodies, the pool was supplemented with the relevant studies in the literature as follows.

a. Normative definition of professionalism by Swick (2000)

Swick (2000) from the University of Montana, USA, has introduced a normative definition for medical professionalism, which consists of nine elements (Table 18), claiming that these elements are 'grounded in what doctors actually do and how they act individually and collectively' (p. 614). This definition is somewhat unique to the other studies of professionalism (Cruess *et al.* 2004; Epstein & Hundert 2002) as the nine elements are in operationalised terms, i.e. immediately transferable to the working life of a doctor.

Table 18- Nine elements of professionalism as described by Swick (2001)

Physicians subordinate their own interests to the interests of others
Physicians adhere to high ethical and moral standards
Physicians respond to societal needs, and their behaviours reflect a social contract with the communities served
Physicians evince core humanistic values, including honesty and integrity, caring and compassion, altruism and empathy, respect for others, and trustworthiness
Physicians exercise accountability for themselves and for their colleagues
Physicians demonstrate a continuing commitment to excellence
Physicians exhibit a commitment to scholarship and to advancing their field
Physicians deal with high levels of complexity and uncertainty
Physicians reflect upon their actions and decisions

O'Sullivan and Toohey (2008) in Australia developed professionalism vignettes based on Swick's definition, which helped them identify the gaps in professionalism among undergraduates, and compare the differences in perception of professionalism between students and faculty. Blue *et al* (2009) evaluated the knowledge and attitudes of US undergraduates in professionalism against Swick's definition. The alignment between the perceptions of students and the definition was found to be moderate. Cottrell *et al* (2006) developed a peer assessment tool for US medical undergraduates based on a professionalism framework underpinned by this definition, which was demonstrated to be effective both educationally and psychometrically. All authors, who used this definition, claimed that it was an acceptable and operational conceptualisation of professionalism. Therefore, the professionalism elements in Swick's definition were considered.

b. Professionalism Mini-Evaluation Exercise (P-MEX) by Cruess *et al* (2006)

As the P-MEX is one of the leading assessment tools developed to date to assess professionalism of individual students in multiple settings (Cruess *et al.* 2006), the 24 attributes in P-MEX were also considered. The stages of assessment development were

followed to a great extent in formulating this tool; a large number of behavioural items were identified from the literature, and the number of items was distilled down to a representative and manageable number with the consensus of experts. The items (Table 19), therefore, do not directly represent or operationalise a professionalism framework of a regulatory or professional body. These behaviours were field-tested and validated in a multi-centre study conducted in Japan and reported to be acceptable to different cultures (Tsugawa *et al.* 2011). The same set of behavioural items was used as a primary source in a measure developed for the assessment of professionalism of staff members, providing further evidence of relevance and validity of the constituent items (Todhunter *et al.* 2011).

Table 19 - Professionalism Mini Evaluation Exercise (P-MEX) - subscales and behaviours (Cruess *et al.* 2006)

Subscales	Behaviours
Doctor–patient relationship skills	Listened actively to patient
	Showed interest in patient as a person
	Showed respect for patient
	Recognized and met patient needs
	Accepted inconvenience to meet patient needs
	Ensured continuity of patient care
	Advocated on behalf of a patient and/or family member
	Maintained appropriate boundaries with patients/colleagues
Reflective skills	Demonstrated awareness of limitations
	Admitted errors/omissions
	Solicited feedback
	Accepted feedback
	Maintained composure in a difficult situation
Time management	Was on time
	Completed tasks in a reliable fashion
	Was available to patients or colleagues
Interprofessional relationship skills	Maintained appropriate boundaries with patients/colleagues
	Maintained appropriate appearance
	Addressed own gaps in knowledge and skills
	Demonstrated respect for colleagues
	Avoided derogatory language
	Assisted a colleague as needed
	Maintained patient confidentiality
	Used health resources appropriately
	Respected rules and procedures of the system

c. Two studies focused on junior undergraduate level

Two studies, which focused on the early stages of medical undergraduate education, were considered, as the literature suggested the existence of stages of professionalism development among students, i.e. certain professionalism attributes of doctors may not make sense to junior medical students (Hilton & Slotnick 2005). The first study by Maudsley *et al* (2007) established the characteristics of a good doctor through the

perceptions of junior medical students while the second study by Papadakis *et al* (2001) proposed an assessment form focused on the professional behaviour first and the second year medical students.

Expectations of medical students of a good doctor by Maudsley et al (2007)

This study sought the perceptions of nearly 1000 junior medical students in a UK medical school about their expectations of a 'good doctor' using a mixed method approach (Maudsley *et al.* 2007). The study participants belonged to two entry cohorts and the authors achieved a 91% response rate. The findings should be credible as both the number of responses and response rate were high and the methodology was rigorous. As the study was conducted in the UK undergraduate setting and represent the students' perspective. Therefore, the findings appeared to be more relevant to this doctoral project. Therefore, the elements identified in this study (Table 20) were included.

Table 20 - Characteristics of a good doctor as identified by UK medical students (Maudsley *et al.* 2007)

Themes of a 'good doctor'	Examples
Compassionate, patient-centred carer	Caring, empathetic, supportive, compassionate, considerate)
Listening, informative communicator	listens well, communicates well, is willing/able to explain in an understandable way
Exemplary, responsible 'professional'	Assertive, responsible, trustworthy, dignified, resilient, professional, hardworking, tactful
Experienced, knowledgeable expert	Knowledgeable, knows the job, intelligent, experienced, academically able, educated
Friendly, inclusive team player	A team player, friendly, non-judgemental, approachable
Thinking, flexible learner	Up-to-date, aware of own limits/fallibility, quick thinking, open-minded, curious, creative
Decisive, competent diagnostician	Good at diagnosing/treating, competent, well trained
Well balanced, insightful 'individual'	Happy, self-aware, work-play balanced, optimistic, humorous
Efficient, organised self-manager	Efficient, organised, good at time-management

Assessment of professionalism among first and second year students by Papadakis et al (2001)

Although the assessment form was developed in a US medical school by academic staff, the behaviours identified in this form (Table 21) were based on the common complaints of unprofessionalism among junior medical students. However, these behaviours were assessed longitudinally throughout the undergraduate curriculum of this medical school. Therefore, the authors believed that a common set of behaviours could be used for the assessment of professionalism across different levels of study.

Table 21 - Professionalism assessment form for first and second year medical students

Aspect of professionalism	Attributes
Reliability and responsibility	Fulfilling responsibilities in a reliable manner.
	Learning how to complete assigned tasks.
Self-improvement and adaptability	Accepting constructive feedback
	Recognizing limitations and seeking help
	Being respectful of colleagues and patients
	Incorporating feedback in order to make changes in behaviour
	Adapting to change
Relationships with students, faculty, staff and patients	Establishing rapport
	Being sensitive to the needs of patients
	Establishing and maintaining appropriate boundaries in work and learning situations
	Relating well to fellow students in a learning environment
	Relating well to staff in a learning environment
	Relating well to faculty in a learning environment
Upholding the Medical Student Statement of Principles	Maintaining honesty
	Contributing to an atmosphere conducive to learning
	Respecting the diversity of race, gender, religion, sexual orientation, age, disability or socioeconomic status
	Resolving conflicts in a manner that respects the dignity of every person involved
	Using professional language and being mindful of the environment
	Protecting patient confidentiality
	Dressing in a professional manner

d. Professional attributes of General Practitioners (van de Camp *et al.* 2006)

Some evidence suggests that the focus on certain aspects of professionalism could vary between primary care and hospital-based practices due to the differences in setting, patient expectations and social image (Natanzon *et al.* 2010; Williams & Calnan 2002). Therefore, the behaviours included in an assessment instrument of professionalism in a general practice setting in the Netherlands (van de Camp *et al.* 2006) were considered. (Table 22)

Table 22 - Professional attributes of the General Practitioner (van de Camp *et al.* 2006)

Towards patients	Towards other professionals	Towards the public	Towards oneself
Integrity	Transmural care	Accountability	Self-reflection
Detachment and commitment	Co-operation with specialists	Ability to make use of the opportunities of the profession	Self-confidence
Respect	Co-operation with support personnel	Norms and values	Self-welfare
Dealing with patient diversity	Leadership	Quality management	Provide and receive feedback
	Collegiality	Practice management	Life-long learning
			Resilience
			Dealing with mistakes
			Dealing with uncertainty
			Cope with aggression
			Resilience

3.2.2. Dealing with 'grey' areas

In the review of the literature, certain areas were found to be complex and at times, it was not easy to categorise or group the attributes. These areas with less clarity are discussed below.

'Altruism' has been defined in different ways, making its interpretation difficult. For example, the ABIM framework defines altruism as the prioritising the interests of patients over the doctor's own interests (ABIM 2001). However, in CanMEDS roles altruism is considered as commitment (RCPSC 2005). Although the ABIM interpretation of altruism was quite similar and more acceptable to the UK understanding of the term, the concept itself has been debated in the UK (RCP 2005), with the resultant definition of 'unselfish approach to patient care' (RCP 2005).

'Excellence' in the UK context is maintaining the highest standards of practice (RCP 2005b). However, distinguishing between this concept and 'mastery of the discipline' (RCPSC 2005) or 'clinical competence' (ABIM 2001) used by other conceptual frameworks was found to be difficult. It appears to overlap with terms such as 'keeping up-to-date' and 'lifelong learning'. Therefore, it was decided to abandon the term 'excellence' and present the ideology behind it more explicitly with behaviours like 'maintaining standards in patient care', 'commitment to improvements' and 'reflective practice'.

'Moral contract' has been introduced as an attribute of professionalism after amalgamating morality and social contract (RCP 2005). Moral contract, therefore, is the behavioural complex represented by personal qualities such as honesty and integrity, accountability and behaving responsibly towards work and society (Cruess *et al.* 2000a).

Table 23 summarises the professional attributes identified from the sources discussed above.

Table 23 - Attributes of professionalism as identified in the literature

	GMC (GMC 2009)	RCP (RCP 2005)	ABIM (ABIM 2001)	RCPSC (RCPSC 2005)	Swick (2000)	PMEX (Cruess *et al.* 2006)	Papadakis (2001)	van de Camp (2006)	Empathy scale (Hojat *et al.* 2001)	Lifelong learning scale (Hojat *et al.* 2003b)	Polyporfessionalism Inventory (Roff *et al.* 2011)	Maudsley (2007)
1. Having the necessary clinical competence	x	x		x		x	x					X
2. Being respectful towards patients', colleagues', other professionals' and students	x	x	x	x	x	x	x	x				X
3. Practicing ethically (maintaining confidentiality and privacy, and obtaining consent when and where necessary	x	x	x	x	x	x	x					X
4. Being responsible towards work and society	x	x	x	x	x		x					X
5. Behaving in a reliable and dependable way	x					x	x					
6. Being accountable for one's actions	x	x	x	x	x	x		x				
7. Showing compassion towards patients	x	x		x	x	x						x
8. Having a caring attitude towards patients	x	x			x	x						x
9. Being truthful in verbal and written communications	x		x		x							x
10. Maintaining standards in patient care and academic work (including research)	x	x		x	x	x		x				x
11. Keeping up-to-date	x	x		x		x						x
12. Practising self-directed learning	x			x								
13. Commitment to improvement	x	x	x	x	x							
14. Being helpful to patients, colleagues and other healthcare professionals	x		x			x	x	x				x
15. Being accessible to patents and other healthcare professionals	x		x			x						x
16. Conforming to rules, regulations and professional standards	x		x	x		x		x				
17. Non-exploitation of position to take personal professional advantages	x		x									
18. Being sensitive to the cultural background of colleagues and patients	x		x	x			x					

	GMC (GMC 2009)	RCP (RCP 2005b)	ABIM (ABIM 2001)	RCPSC (RCPSC 2005)	Swick (2000)	PMEX (Cruess *et al.* 2006)	Papadakis (2001)	van de Camp (2006)	Empathy scale (Hojat *et al.* 2001)	Lifelong learning scale (Hojat *et al.* 2003b)	Polyporf Inventory (Roff *et al.* 2011)	Maudsley (2007)
19. Appreciation of diversity specially gender, age, race or ethnicity	x		x				x	x				x
20. Being empathetic when caring for patients	x		x		x				x			x
21. Managing conflicts of interest	x		x	x	x		x					
22. Being aware of one's limitations as a practitioner	x		x			x	x					x
23. Ability to work in a team	x	x	x				x					x
24. Using resources appropriately	x					x						
25. Consider appropriate appearance, dress and personal hygiene	x					x						
26. Reflection 'on' and 'in' practice	x	x		x	x			x				x
27. Commitment to life-long learning	x	x	x	x	x	x		x		x		
28. Prudence in making decisions	x	x			x			x				x
29. Attention to self-care	x			x				x				x
30. Providing advice to patients and colleagues when required	x		x	x		x						
31. Participation in social, institutional and personal development	x	x			x		x					
32. Involving in teaching, training and supervision	x			x								
33. Acting with confidence in one's duties	x							x				x
34. Maintaining professional boundaries	x					x		x				
35. Ensuring continuity of care	x											
36. Evidence-based practice	x		x									
37. 'Whistle-blowing '	x	x	x									
38. Being adaptable to changes in the workplace	x						x					
39. Being honest and acting with integrity in academic, clinical administrative work	x	x	x	x	x		x	x			x	
40. Being collegiate in workplace	x		x					x				
41. Showing altruism towards patients		x	x	x	x		x					
42. Avoiding substance or alcohol misuse											x	
43. Avoiding a cynical approach in one's job		x	x		x							

	GMC (GMC 2009)	RCP (RCP 2005b)	ABIM (ABIM 2001)	RCPSC (RCPSC 2005)	Swick (2000)	PMEX (Cruess *et al.* 2006)	Papadakis (2001)	van de Camp (2006)	Empathy scale (Hojat *et al.* 2001)	Lifelong learning scale (Hojat *et al.* 2003b)	Polyporf Inventory (Roff *et al.* 2011)	Maudsley (2007)
44. Considerate on social justice		x										
45. Ability to collaborate		x										
46. Conforming to both social and professional etiquette		x			x			x				
47. Having leadership skills		x						x				
48. Showing equity in providing opportunities			x									x
49. Maintaining composure / Equanimity especially in difficult situations			x			x		x				
50. Prioritising service interests over commercial interests			x			x	x					
51. Being punctual						x						
52. Avoiding deception			x									
53. Dedication towards professional work			x		x							
54. Being receptive to constructive criticism			x			x	x	x				
55. Absence of impairments			x		x							
56. Communicating in a clear and effective manner						x	x					x
57. Being resilient								x				x
58. Having skills in practice management								x				

3.2.3. Ensuring the clarity and relevance of items

Clarity, specificity and conciseness of items determine the properties and utility of a measure (DeVellis 2003, p.64). Many researchers (Blackall *et al.* 2007; Hojat *et al.* 2001; Todhunter *et al.* 2011) used the opinion of a large group of experts in the development of items to ensure clarity and relevance. Thus, the opinion of an international group of 26 medical educators (Table 24), who participated in a face-to-face session on professionalism at the Centre for Medical Education, University of Dundee in April 2009, was sought. An international group was preferred to a group of UK nationals alone, as the clarity and meaning of the terms used in the measure will be acceptable to a wider population in event of validating the final measure in countries other than the UK.

Table 24 - The geographical origin of medical educationalists who participated in the item refining exercise

	Frequency	%
Middle East	7	26.92
Asia	1	3.85
Africa	3	11.54
Australasia	1	3.85
Europe (UK)	9 (5)	34.62
North America	5	19.23
Total	26	100.00

They reviewed the items (attributes) for clarity, relevance and mutual exclusiveness during a session that lasted one hour. In the discussion, the opinions of all members were considered equally in determining the clarity, wording and mutual exclusiveness. However, when there was a conflict of opinions, the priority was given to the opinions of the UK participants in determining the relevance as the final measure would focus primarily on the UK context. This resulted in 46 items (Table 25).

Table 25 - Revised list of items after the feedback session with a group of medical educators

(The 58 items in this table is put in a different order to Table 23 above to group the select, altered and non-selected items together.)

Items developed by the author	The final form of items selected to proceed with	Important issues rained / observed in the discussion which helped take decisions
Decided to include without change		
1. Showing altruism towards patients	1. Showing altruism towards patients	
2. Communicating in a clear and effective manner	2. Communicating in a clear and effective manner	
3. Behaving in a reliable and dependable way	3. Behaving in a reliable and dependable way	
4. Avoiding substance or alcohol misuse	4. Avoiding substance or alcohol misuse	Although this was identified only in a single measure participants were of the view that it represents behaviour of doctors as a role model in society.
5. Being accountable for one's actions	5. Being accountable for one's actions	
6. Avoiding a cynical approach in one's job	6. Avoiding a cynical approach in one's job	In their experience, the participants had identified a cynical approach to the job in newer generations of doctors
7. Being sensitive to the cultural background of colleagues and patients	7. Being sensitive to the cultural background of colleagues and patients	
8. Being empathetic when caring for patients	8. Being empathetic when caring for patients	
9. Being aware of one's limitations as a practitioner	9. Being aware of one's limitations as a practitioner	
10. Being punctual	10. Being punctual	
Decided to include after rephrasing		
11. Non-exploitation of position to take personal professional advantages	11. Not using one's professional status for personal gain	
12. Maintaining composure / Equanimity especially in difficult situations	12. Behaving with composure	

13. Using resources appropriately	13. Making effective use of the resources available	
14. Considering appropriate appearance, dress and personal hygiene	14. Being mindful of one's personal appearance	
15. Conforming to both social and professional etiquette	15. Conforming to social norms	
16. Dedication towards professional work	16. Taking a dedicated approach to one's work	
17. Prudence in making decisions	17. Being sound in judgment and in decision making	
18. Attention to self-care	18. Looking after one's own health and well-being	
19. Providing advice to patients and colleagues when required	19. Providing advice to patients and colleagues when required	
20. Being receptive to constructive criticism	20. Being receptive to constructive criticism	
21. Having leadership skills	21. Showing leadership skills and initiative	
22. Acting with confidence in one's duties	22. Acting with confidence in one's duties	
23. Being adaptable to changes in the workplace	23. Being adaptable to changes in the workplace	
24. Involving in teaching, training and supervision	24. Having the skills to train colleagues if required	
Decided to combine / split		
25. Being respectful towards patients', colleagues', other professionals' and students	25. Respecting patients autonomy 26. Respecting patient's confidentiality and privacy 27. Respecting colleagues	The participants were of the view that 'practising ethically' was too broad and non-specific, and the items should be presented more explicitly. Some expressed the view that ethics here should reflect the principles of medical ethics { autonomy, beneficence, non-maleficence, and justice (Gillon 1994)}. As beneficence and maleficence overlap largely with several other items included in the set of items it was decided to include the principles of 'patient autonomy' and, 'maintaining confidentiality and privacy'. It was argued that, in practice, some doctors do not respect patients but respect their colleagues; therefore respecting patients and colleagues should be separate.
26. Practising ethically (maintaining confidentiality and privacy, and obtaining consent when and where necessary		

27. Being responsible towards work and society	28. Acting in a responsible fashion towards patients 29. Acting in a responsible fashion towards colleagues 30. Acting in a responsible fashion towards society	The participants were of the view that doctors' responsibility towards work ultimately affects patients and colleagues. Therefore, the 'responsibility' should reflect patients, colleagues (including students) and society.
28. Showing compassion towards patients	31. Showing compassion towards one's patients	It was expressed that dividing 'caring' and 'compassion' is difficult and may not be necessary. It was anticipated that these two items would not be effective discriminators individually.
29. Having a caring attitude towards patients		
30. Being truthful in verbal and written communications	32. Behaving honestly and with integrity	The participants suggested that 'being truthful' and 'avoiding deception' are also related to honesty and integrity
31. Being honest and acting with integrity in academic, clinical administrative work		
32. Avoiding deception		
33. Maintaining standards in patient care and academic work (including research)	33. Having a positive attitude towards professional development	The majority expressed that keeping up-to-date, practising self-directed learning and commitment to improvement are related to a positive attitude towards professional development. However, the opinion was divided over the suggestion of including the maintenance of standards of care and academic work under the professional development umbrella. Eventually agreement was reached to include it under 'professional development'.
34. Keeping up-to-date		
35. Evidence-based practice		
36. Practising self-directed learning		
37. Commitment to improvement		
38. Being helpful to patients, colleagues and other healthcare professionals	34. Being attentive to the needs of colleagues 35. Being attentive to the needs of patients	The participants perceived that patients and colleagues in this item should be kept separate due to the same reasons expressed under items 26 and 27 above.
39. Being accessible to patents and other healthcare professionals	36. Being accessible to patients 37. Being accessible to colleagues	

40. Conforming to rules, regulations and professional standards	38. Adhering to professional rules and regulations 39. Functioning according to the law	The participants suggested that 'functioning according to the law' should be a separate item. Some of them of the view that doctors at time have certain level of privilege in the law enforcement. As having a privileged status was an element in the 'old' definition of professionalism but have no validity in the present context (Wass 2006), it was decided to include it as a separate item.
41. Appreciation of diversity specially gender, age, race or ethnicity 42. Showing equity in providing opportunities 43. Considerate on social justice	40. Treating patients fairly and without prejudice 41. Treating other healthcare professionals fairly and without prejudice 42. Treating colleagues fairly and without prejudice	The participants perceived that the appreciation of diversity, equity and social justice are about being fair and not being prejudiced. They were keen on considering other healthcare professionals separately as hierarchy and prejudice may be more evident within multi-professional teams.
44. Prioritising service interests over commercial interests 45. Managing conflicts of interest	43. Being able to manage situations where there is a conflict of interest	It was argued that both items were focused on managing conflicting interests.
46. Ability to work in a team 47. Being collegiate in workplace	44. Working well as a member of a team	'Collegiality' was misunderstood by many including the native English speakers in the group. The participants were not convinced that it would be easily understood by medical students either. Working in a team was considered a better way of explaining this concept.
48. Reflection 'on' and 'in' practice 49. Commitment to life-long learning	45. Reflecting on one's actions with a view to improvement	Reflection was seen as the basis of lifelong learning and thus it was suggested to combine the two items.
50. Ability to collaborate 51. Participation in social, and institutional development	46. Working with one's colleagues towards common goals	Both collaboration and participation in institutional developments were considered by many participants as working with others in the healthcare team to achieve common goals.
Decided to exclude		
52. Having the necessary clinical competence	Excluded	Including clinical competence in the inventory was discussed in length. It was argued that this item largely represents knowledge and technical skills and therefore, should not be included. However, opponents argued that it is a key component of several professionalism frameworks. Based on the nature of this measure, i.e. the professionalism culture of an institution rather than the professionalism of individuals,, it was decided to exclude the item from the list. Respondents, e.g. students, may not be

		able to comment of the overall clinical competence of clinical teachers. The literature also suggests that clinical competence is not necessarily an attribute of professionalism (Todhunter *et al.* 2011).
53. Being resilient	Excluded	The participants were not certain about the actual meaning of the item and suggested exclusion.
54. Having skills in practice management	Excluded	It was perceived that this item is highly specific to the general practice setting.
55. Absence of impairment	Excluded	As personal impairments are not always in the public domain it was perceived that the respondents to the final survey may be unable to respond to this in a meaningful way.
56. Maintaining professional boundaries	Excluded	The participants were of the opinion that this may be misinterpreted as 'professional autonomy' or 'hierarchy'. As the intended meaning of this item is addressed with items related to ethics, it was suggested to exclude this item.
57. Ensuring continuity of care	Excluded	It was perceived that this item is represented by many other items, e.g. responsible behaviour, team work.
58. 'Whistle blowing'	Excluded	The participants argued that this is part of accountability.

3.3. Discussion

On examination of the items, it is evident that they do not demonstrate the key weaknesses associated with 'bad' items such as being unusually lengthy or being difficult or ambiguous to understand (Clark & Watson 1995; DeVellis 2003, pp.67 - 69). The items list, however, contains a very few items, in which more than one aspect needs to be considered in responding, i.e. double barrel items (e.g. being sensitive to the cultural background of *colleagues* and *patients*). Although this may cause difficulties for respondents (Clark & Watson 1995; DeVellis 2003, pp.67 - 69), the group of reviewers felt strongly that it would have no adverse effect on the measure.

These attributes overlap considerably with other lists identified in the professionalism literature. For example, Todhunter *et al* (2011) developed a tool to assess the professionalism of faculty in the Canadian context. Their 35 attributes were very similar to those identified in the present study. Todhunter *et al* (2011) excluded 'knowledge' and 'competence', 'applying ethical principles', and 'ensuring continuity of care'. However, unlike in the present study Todhunter *et al* (2011) did not include 'patient advocacy', 'enthusiasm towards teaching', 'the ability to resolve conflicts', 'the achievement of goals collaboratively', 'maintaining composure', 'punctuality', and 'the use resources appropriately'. Again, the themes identified by van de Camp *et al* (2004) in their literature review were similar to the present study. However, they included 'self-regulation', 'absence of impairments' 'submission to ethics code', 'excellence', 'social contract' and 'response to stress'. 'Self-regulation', however, is an attribute of 'old-professionalism' that

has been replaced by public scrutiny in the 'neo-professionalism' movement (Wass 2006). 'Absence of impairments' is an independent fitness-to-practice issue which is separate from professionalism (Parker 2006). Broad areas such as 'ethical conduct', 'excellence' and 'social contract' are indirectly covered by other items in the present study. According to van de Camp *et al* (2004), adherence to professional guidelines, self-improvement, humanistic values, avoidance of prejudice and discrimination, patient autonomy, which are attributes of professionalism identified in the present study, were less popular in the literature. Since 2004, when van de Camp *et al* published this article, the constituents of professionalism have transformed considerably. In the review of the literature by Wilkinson *et al* (2009), who followed a similar approach to van de Camp *et al* (2004), the attributes of professionalism identified were largely in concordance with the present study. Common themes between the review by Wilkinson *et al* (2009) and the present study included honesty and integrity, confidentiality, moral reasoning and respecting professionals' privileges. However, 'maintaining professional boundaries' and 'an organised approach to work' did not appear in the present study.

This process may have several limitations. In the development of a measure for the UK, the use of an international group of medical educators may have caused a certain amount of bias. However, the process was helpful to enhance the acceptability and comprehensibility of the items, which would be included in the final measure, is acceptable to a wider population. To minimise the bias on the relevance of items to the UK context, preferences was given to the opinion of the UK group members in

determining the relevance. The interpretation of items during the process of refinement may have been influenced by the author's background. The author was born and trained in eastern part of the world where the interpretation of professionalism may be totally different to the west (Ho *et al.* 2012). To minimise this bias, the author was in a regular dialogue with the UK medical trainees and practitioners to achieve near-accurate interpretations for different professional attributes.

In conclusion, although there are minor differences, the list of attributes identified for use in the validation study was largely compatible with the published reviews of the literature thus supporting the credibility of the refinement process. Minor differences between reviews could be due to differences in criteria setting and chronology as professionalism attributes do appear to change over time. The refined set of attributes was used for the validation surveys conducted among the UK general public and medical educators.

Chapter 4 – Achieving consensus of stakeholders

4.1. Introduction and justification of the methodological approach

In establishing content validity of a psychological measure achieving consensus of experts on the relevance and importance of its items is a primary requirement (Johnston & Wilkinson 2009; Lawshe 1975; Polit & Beck 2006). Not all measures of professionalism culture (Table 16, p.1) have followed such a thorough process prior to field-testing (Arnold *et al.* 1998; Kalet & Steven 2004; Quaintance *et al.* 2008; Roff *et al.* 2011; Thrush *et al.* 2011; Wiggleton *et al.* 2010). However, the rigour of the validation process determines the robustness of the final measure and it is recommended as an essential stage as well as a good practice (Clark & Watson 1995; DeVellis 2003, pp.85 - 87). This stage has been followed by many who developed measures of professionalism culture (Blackall *et al.* 2007; Hojat *et al.* 2001; Hojat *et al.* 2003b; O'Sullivan & Toohey 2008; Roff *et al.* 1997).

As professionalism is increasingly recognised as a social contract between doctors and society (Cruess & Cruess 2000; Wass 2006), both need to be considered as essential stakeholders in the validation process. The general public may not be considered as 'experts' on the topic, but their perspectives should be duly considered as the conceptualisation of professionalism among the two groups may not necessarily be the same. For example, a qualitative focus group study conducted with postgraduate trainee doctors and members of patients' families in North Carolina reported that doctors prioritised tactfulness and support of team members, while the families emphasised knowledge and skill, and honesty (Regis *et al.* 2011). In a literature review, Hutchings and

Rapport (2012) conclude that even the understanding of patient-centeredness, the essence of modern-day professionalism (Irvine 2001), between practitioners and the public appears to be somewhat incongruent. However, these studies also suggest that there are common grounds between the two groups (Hutchings & Rapport 2012; Regis *et al.* 2011).

The Delphi process and Nominal Group method are the most commonly used methods of achieving consensus for research purposes (Clayton 1997; Jones & Hunter 1995). The Delphi is an iterative process in which a panel of experts ranks a pool of items individually and independently according to their perceived importance in the first round, and change /retain their priorities based on the responses of others, which are fed back to the panel in subsequent rounds (Clayton 1997; Jones & Hunter 1995; Rowe & Wright 1999). The experts can be a purposefully selected, homogenous (e.g. clinicians) or heterogeneous (e.g. clinicians and patients) group (Jones & Hunter 1995; Powell 2003). The panel usually consists of between 10 and 30 experts. The communication between them and the researcher(s) is usually via postal or online questionnaires, and the panellists do not assemble together during the process. The steps frequently followed in a Delphi process are shown in Table 26.

Table 26 - Steps followed in a conventional Delphi process (after Jones & Hunter 1995, p.376)

1. Round 1: Either the relevant individuals are invited to provide opinions on a specific matter, based on their knowledge and experience, or the team undertaking the Delphi expresses opinions on a specific matter and selects suitable experts to participate in subsequent questionnaire rounds
2. These opinions are grouped together under a limited number of headings and statements drafted for circulation to all participants on a questionnaire
3. Round 2: Participants rank their agreement with each statement in the questionnaire
4. The rankings are summarised and included in a repeat version of the questionnaire
5. Round 3: Participants re-rank their agreement with each statement in the questionnaire, with the opportunity to change their score in view of the group's response
6. The re-rankings are summarised and assessed for degree of consensus: if an acceptable degree of consensus is obtained the process may cease, with final results fed back to participants; if not, the third round is repeated

The Nominal Group technique is a highly structured face-to-face exercise in which the information about a given topic is gathered from a small number of experts (usually 9 – 12) (Jones & Hunter 1995). The steps of a Nominal Group meeting are shown in Table 27.

Table 27- Steps followed in a typical Nominal Group technique (after Jones & Hunter 1995, p.376)

1. Participants spend several minutes writing down their views about the topic in question
2. Each participant, in turn, contributes one idea to the facilitator, who records it on a flip chart
3. Similar suggestions are grouped together, where appropriate. There is a group discussion to clarify and evaluate each idea
4. Each participant privately ranks each idea (round 1)
5. The ranking is tabulated and presented
6. The overall ranking is discussed and reranked (round 2)
7. The final rankings are tabulated and the results fed back to participants

The Delphi process is the most commonly used method of achieving consensus in establishing content validity of measures on professionalism culture to date (Blackall *et al.* 2007; Hojat *et al.* 2001; Hojat *et al.* 2003b; O'Sullivan & Toohey 2008; Roff *et al.* 1997). van de Camp *et al* (2006), established the content validity of an instrument, which assessed

the professional behaviour of general practice trainees in the Netherlands, using the Nominal Group method.

The Delphi process makes allowance for opinions of a large number of participants compared to the Nominal Group technique, which is restricted by the necessity of gathering the participants together on a particular day (Clayton 1997). The independent and individualised nature of the Delphi process helps reduce the monopolisation of group discussions by dominant individuals and increases the orientation of the discussion towards the problem (e.g. what is professionalism?) rather than the interests of a group (e.g. how our opinions on professionalism as clinicians may affect medical practice) which may happen in Nominal Group meetings (Clayton 1997). In a Delphi panel, the smaller the number the less the generalisabilty of results and the larger the number the lower the response rate / higher the attrition (De Villiers *et al.* 2005). The researcher should be familiar with the experts on the panel, and should select those with a high degree of ego involvement, who may less likely to drop out or abandon the Delphi process (Bardecki 1984).

Although the Delphi process could have been considered for this study the weaknesses associated with its conventional use may have several negative implications. Having both practitioners and members of the general public may lead to an undesired and unreliable outcome. Although professionals are eligible to serve in a Delphi panel (De Villiers *et al.* 2005), obtaining the consent of an adequate number and retaining them throughout the

process would be a challenge, particularly in a field such as medicine where many practitioners are restricted by time pressures.

The conventional version of Delphi, therefore, did not appear to be appropriate for the current study. It was decided to use a survey method with the view to accessing a large and representative sample of the general public and professionals, and to statistically analyse the results to establish content validity (Johnston & Wilkinson 2009; Lawshe 1975; Polit & Beck 2006; Rubio *et al.* 2003).

In a consensus process, the agreement should be described not only in relation to the extent to which each respondent agrees with the issue under consideration, but also in relation to the extent to which respondents agree with each other (Jones & Hunter 1995); the consensus within each group and between the two groups, therefore, had to be determined.

The specific objectives of the surveys were:

a) to determine the 'public model of professionalism' based on the importance placed on the professional attributes by the UK general public;
b) to identify the essential attributes of professionalism as perceived by medical professionals in the UK with acceptable consensus; and
c) to select the items, which are important to both the general public and practitioners, to represent the public model of professionalism identified in objective (a).

4.2. Methodology

4.2.1. Ethics

The ethics permission to this project was granted by the University Research Ethics Committee (UREC), University of Dundee, UK [Reference Numbers: UREC 9026 (Public survey- Appendix I) and UREC 10047 (Professionals' survey – Appendix II)].

4.2.2. Study design

In order to achieve the above specific objectives, two cross-sectional quantitative studies were conducted sequentially among the general public and medical professionals (educationalists and practitioners) in the UK. The aim of both studies was to assess the perceived importance of the 46 professional attributes in order to retain the most important / essential attributes.

4.2.3. The survey

The questionnaires survey consisted of 55 items with a five-point rating scale.

a. The survey items

All 46 attributes of professionalism identified in Chapter 3 are socially desirable and morally appealing. Therefore, it was anticipated that the respondents would blindly rate them with equal importance, resulting in non-discriminatory results; such responses can be a potential drawback in using the survey method (Meade & Craig 2012). The use of a mixture of positively and negatively worded items is popular in minimising this drawback. However, as discussed in the literature review (p.1), this approach may not alleviate the problem, but complicate it further (DeVellis 2003, pp. 69 & 70; Meade & Craig 2012;

Schuman & Presser 1996, pp. 203 - 230). Alternatively, there are several methods suggested to detect inattentive responding prior to the administration of surveys, e.g. inclusion of lie scale, bogus items, instructed response items (Meade & Craig 2012). These items are expected to be rated and scored differently to the actual items (Hargittai 2009). Bogus items have been used in several healthcare surveys to discriminate between attentive and non-attentive responding to items (Deshpande 2009; Green *et al.* 2009; Soriano *et al.* 2011). As this strategy has been very effective in the contexts of both healthcare professionals and the public, the inclusion of bogus items was considered appropriate for this survey. Accordingly, to enhance the credibility of responses, nine societally perceived attributes of professionalism based on the author's own experience and the experiences of his supervisors, but not evidence-based, were included in the survey (Table 28). This complied with the need to use bogus items, which are related to the measured construct, but not evidence based (Hargittai 2009). Therefore, the final validation surveys consisted of a total of 55 items; 46 were evidence-based and nine were not.

Table 28 - Non-evidence based items included in the public and professionals' surveys

Having a good sense of humour
Being physically fit
Being well read outside the professional arena
Always being busy
Earning a high salary
Speaking with a refined accent
Attending a prestigious school
Being physically attractive
Owning a luxurious home

In addition to non-attentive responding, the sequence in which items appear in the survey may lead to two response errors. The first, proximity error, which occurs when items appear to be related but distinct, are placed close to each other; the response to one is influenced by the other. For example, if 'being accessible to patients' is placed just after 'being accessible to colleagues', the response to one may be influenced by the response to the other; respondents may not think explicitly about patients and colleagues. Closely related items were, therefore, placed well apart as a countermeasure. (McAleer & Chandratilake In press; Russell 1979). The second, logical error occurs when the response to one item influences the response to another despite placing them widely apart (McAleer & Chandratilake In press). The enhanced clarity of items is thought to be helpful in minimising this error. Further measures were taken at the delivery stage to reduce the effects of these errors, which are discussed below in the appropriate sections.

b. The rating scale

As ranking 55 items was impractical, the only possibility was to request the respondents to select a given number of items (e.g. 20) in order of their preference. In obtaining perceptions, both rankings and rating scales have performed equally in terms of validity and reliability at least in undergraduate settings (Rankin & Grube 1980). However, unlike rating scales, rankings, which involve forced-choice, i.e. ipsative data (Cornwell & Dunlap 1994), cannot be subjected to advanced statistics such as factor analysis (Cornwell & Dunlap 1994). Therefore, for these surveys, the rating of each item based on perceived importance was preferred to the selection of items by ranking.

It was anticipated that all 46 items would have some importance as they originated mainly from the literature. Therefore, a unipolar scale, i.e. more categories for either positive or negative side (more categories on the positive side in this instance) was preferred to a bi-polar scale, i.e. equal number of categories for both positive and negative sides, for the purpose of this survey (DeVellis 2003, pp74-85; Green *et al.* 2009). Having more than four categories in the rating scale has been shown to yield responses which would reliably discriminate between items (Lozano *et al.* 2008) and the mean score obtained through a five-point scale would not be significantly different to a nine-point scale (Cowley & Youngblood 2009). Therefore, a unipolar five-point scale (1 = not important, 2 = slightly important, 3 = somewhat important, 4 = very important, 5 = extremely important) was used in both surveys. The type of scale used should be appropriate to the subsequent analysis of responses (DeVellis 2003, pp74-85). As the method of analysis used in the two surveys is different, how this scale was used in the analysis is discussed below under each survey. A tailored approach to data collection and analysis helps accommodate the differences in intellectual abilities and skills, knowledge of the subject and, the interest in the topic, thereby enhancing the credibility of information gathered (DeVellis 2003).

4.3. The public survey

The objective of the survey was to conceptualise professionalism from the perspective of the UK general public. Therefore, the question posed was: How important are these attributes to you as qualities of your doctor?

4.3.1. Sampling and recruitment

Targeting patients, visiting a particular hospital, clinic or general practice, has been the popular method of acquiring 'public' perceptions on professionalism (Davis *et al.* 2007; Wiggins *et al.* 2009). The results yielded from such a single centre/context approach, however, may not be representative. A national random sample could effectively enhance the representativeness provided that the response rate is high (Green *et al.* 2009). However, patients, a selective group of the public, may not necessarily share the same view as the general public (Slevin *et al.* 1990). The judgements and perceptions of patients on issues related to their doctors may be influenced by the need of maintaining a 'good' relationship and continuity of care (Ginsburg *et al.* 2000). Therefore, accessing a sample of the UK general public, which is representative of the national demographic characteristics, should more accurately describe the public conceptualisation of professionalism. Based on a meta-analysis, it has been estimated that the average response rate of a random-sample of the UK general public to a survey on healthcare-related matters is 53.8% (Baruch & Holtom 2008). When a low response rate is anticipated, ensuring the representativeness becomes vital to maintaining the generalisability of findings (Cook *et al.* 2000), as the non-response bias is always a concern in public surveys (Kellerman & Herold 2001). In fact, it

has been argued that representativeness is more important than the response rate itself in public surveys (Cook *et al.* 2000).

Therefore, it was decided to use the quota sampling technique to achieve a nationally representative sample of the UK general public. Quota sampling is the 'non-probability equivalent of stratified sampling' and helps sample different elements (e.g. age groups, gender, habitat and social class) as they occur in the target population (Cohen & Manion 1994, p.89). A random sample with a low to moderate response rate (which is likely in surveying the public) and a corresponding quota sample produce very similar results; even the significance of the few differences are merely statistical and not practical (Cumming 1990). However, it was practically difficult to achieve a quota sample as the author had no access to personal demographic details of a large group of the general public. Therefore, it was decided to access a section of the general public, who were registered as willing to participate in surveys and research, through a private company (Survey Sampling UK Ltd). The company holds the personal demographic details of their clients with due consent. These clients, however, have not given consent for the researchers to access their individual personal information. Although the quota sampling technique and the use of data bases of the general public held by private agencies are still gaining popularity among researchers working on medical professionalism (Lempp & Seale 2004; Roff *et al.* 2011), similar approaches have been used effectively to obtain the perceptions of the UK general public on many health-related issues, e.g. body-piercing (Bone *et al.* 2008), awareness on cancer screening (Robb *et al.* 2010).

4.3.2. Data collection

The 55-item survey was developed into a web-based questionnaire and the URL link was emailed to the potential participants. As an incentive to participation, £ 0.40 per every completed survey was contributed to a charity fund. Survey Sampling UK Ltd. charged a nominal fee of £0.60 per response to cover administrative expenses. The total cost per respondent, therefore, was one pound. The Centre for Medical Education, University of Dundee, as a part of supporting its PhD students, made the necessary funds available (maximum of £1000) for this survey. Therefore, the target was to achieve a nationally representative quota sample with around 1000 respondents.

Complying with the quota sample technique (Castillo 2009), the age groups, gender and social grades {a computed demographic characteristic, which reflects the occupational and educational level (NRS 2009)} were prioritised as intended quotas. Subsequently, the survey was sent to 1000 members of the general public selected randomly from the pool with a clear explanation of the purpose and uses of the survey. The initial responses were analysed after one week and reminders were sent to non-respondents. Three days after the reminder, more invitations were sent to potential participants, who had not been covered in the first round. This process was repeated until the three quotas were filled. The survey took place between 13 and 29 October 2009. The rationale for the strategies used is discussed below.

a. The web –based format and email delivery

The survey was built into an Adobe Flash® programme by the Survey Sampling UK Ltd. The following features were incorporated.

- Only a single item appeared on the screen at any time. The respondent was expected to drag and drop each item in one of the five boxes, which represented the five options (‘not important’ to ‘extremely important’).
- The subsequent item was accessible only after completing the response to the previous item.
- Accessing previously submitted items, i.e. going backwards, was not possible.
- There was an optional open-comment area at the end which allowed participants to suggest alternative attributes.
- A short evaluation was included as an optional feature to obtaining reaction to the online survey.

The appearance of a single item at a given time and the disallowance of returning to previous answers were thought to be helpful in minimising proximity and logical errors, and in reducing inattentive responding by making the survey interesting. The open comment area helped identify additional areas of professionalism which the public perceived to be important. The evaluation provided an insight into the method used to collect public perception. Once the final technical aspects were completed, the items were reviewed by two public survey advisors in Survey Sampling UK Ltd and they confirmed that the items were comprehensible to all strata of the public. It was subsequently piloted with

14 members of the general public, who reported total satisfaction with of the content and the design.

A meta-analysis of pubic surveys conducted between years 2000 and 2005 demonstrated that the effectiveness (e.g. response rate, quality of responses) of mail and email surveys are not very different if internet access is not a limiting factor (Baruch & Holtom 2008). With 65% of the UK population currently having access to the internet, with impressive coverage across different age, gender and social groups (Office for National Statistics 2008), the argument for online delivery being a possible cause for bias is reduced.

According to the literature, monetary incentives, which are either personally beneficial or charitable, enhance the response rate of online surveys; personal incentives are significantly more effective than charitable incentives (Deutskens *et al.* 2004). Providing personal monetary incentives, however, may influence the sample characteristics, as some tend to complete surveys solely for financial gain (Singer & Couper 2008). Therefore, with the limited funds available for the present project, 40 pence per every completed survey was contributed to a charity fund.

The reminder was timed in compliance with the literature, which suggested that sending a reminder after one week would be most effective (Deutskens *et al.* 2004).

4.3.3. Data analysis

The responses were statistically analysed using Predictive Analytics Soft-Ware (PASW®) version 17 for determining the relative importance of each item, formulating the public model of professionalism and analysing the reaction of respondents to the survey.

a. Determining the importance of items and comparing the responses of demographic groups

The debate over the use of median or mode as the expression of central tendency in analysing parametric data is still unresolved (Carifio & Perla 2008; Jamieson 2004). Theoretically, it is not incorrect to apply parametric statistics on ordinal data provided that the assumptions are clearly stated, and the data is of the appropriate size and shape (Pell 2005). Therefore, it was decided to use either the mean or the mode after analysing the results for normal distribution, i.e. the skewness of the data. The items rated as 'extremely important' and 'very important' important were considered as the 'essential' attributes of professionalism to the general public.

b. Statistically determining the public model of professionalism

Factor analysis helps identify the latent variables, i.e. underlying domains, of a psychological measure and formulate an overall picture of the construct concerned (Clark & Watson 1995; DeVellis 2003, p.102). In the initial analysis of responses, either Exploratory Factor Analysis (EFA) or Principal Component Analysis (PCA) is recommended (Clark & Watson 1995; DeVellis 2003, p.102). Although some advocate that either of these methods can be used indiscriminately on any measure as it makes no different to the final

outcome (DeVellis 2003, p.102), the mathematical model underpinning each method works well for certain designs (Brown 2009b). As EFA uses covariance of item groups, it is more helpful and appropriate in examining a theoretical structure (Brown 2009b). PCA, on the other hand, uses the unique variance of each item to explore all possible patterns emerging from a set of data and it will be more appropriate if a theoretical framework is not used in the development of items (Brown 2009b). However, this advice has not always been adhered to by the developers of psychological measures. For example, in the development of the ABIM Scale of Professionalism, Arnold *et al* (1998) should have used EFA rather than PCA, as the items were developed to represent the ABIM domains of professionalism. Similarly, Cruess *et al* (2006) may be more correct to use EFA instead of PCA in the development of P-MEX, in which the items were not developed with an existing theoretical model in mind. However, PCA was used in the development of Jefferson scales of empathy (Hojat *et al.* 2001) and lifelong learning orientation (Hojat *et al.* 2009a) as they made no presumptions on the theoretical structure for the respective constructs; the advice has, therefore, been upheld.

As there was no consistency in using EFA or PCA in practice, PCA was preferred over EFA in conceptualising the public model of professionalism, complying with the theoretical argument. The group of medical educators made subjective judgements on the items and felt that they are mutually exclusive. Therefore, it was anticipated that there is no-correlation between items; the data were subjected to orthogonal varimax rotation

(Brown 2009a; Hojat *et al.* 2001; Hojat *et al.* 2009a). A factor coefficient of 0.3, the conventional threshold, was considered acceptable for factor loading (Brown 2009a).

c. Reaction of respondents to the survey

The reaction of the respondents' to the survey was evaluated using two questions with Likert-type rating scale and one open comment area. As recommended by Survey Sampling UK Ltd, the two questions represented the standard aspects for the evaluation of participants' experience in public surveys. The data gathered through the Likert-type questions were analysed using descriptive statistics and the open comments were subjected to a thematic analysis.

4.3.4. Results

The results are organised under the following subheadings:

a. Demographic characteristics of the population
b. Importance of each professional attribute
c. Factor analysis

4.3.4.1. Demographic characteristics of the population

It was possible to achieve closely comparable quotas to the UK national population statistics (ONS 2009) in terms of gender, age, social grade and geographical region with 954 respondents (Table 29).

Table 29 - Comparison of the UK national census with the demographic characteristics of public survey respondents

Quota categories	Quotas	Percentage representative of UK general population (ONS 2009)	Respondents achieved in each quota	
			Percentage	Number
Gender	Females	52%	50%	479
	Males	48%	50%	474
Age	18-24	11%	13%	120
	25-44	37%	31%	299
	45-64	31%	32%	301
	65+	21%	24%	233
Social grade	A	11%	16%	150
	B	10%	13%	121
	C	33%	33%	317
	D	26%	20%	194
	E	20%	18%	171
Geographical region	England	84%	86%	816
	Northern Ireland	3%	2%	18
	Scotland	8%	9%	90
	Wales	5%	3%	30

4.3.4.2. Importance of each professional attribute

The responses to all items demonstrated severe to moderate skewness (>1 to 0.5). Therefore, the data was analysed using non-parametric statistics. Accordingly the median values of items indicated that 44 out of the 55 items were essential attributes (rated as either ‘very’ or ‘extremely’ important) of professionalism to the general public. All nine non-evidence-based items together with two evidence-based items (conforming to social norms, being mindful of one’s personal appearance) were rated by the public as ‘not important’ to ‘somewhat important’. The power ratings of all non-evidence based items confirmed the attentiveness of responding (Hargittai 2009). (Table 30)

Table 30 - Items in descending order of importance as perceived by the public

Items with median rating of 5

- Respecting patients autonomy
- Behaving honestly and with integrity
- Respecting patient's confidentiality and privacy
- Acting in a responsible fashion towards patients
- Adhering to professional rules and regulations
- Being attentive to the needs of patients
- Treating patients fairly and without prejudice
- Avoiding substance or alcohol misuse

Items with median rating of 4

- Being accountable for one's actions
- Respecting colleagues
- Communicating in a clear and effective manner
- Acting in a responsible fashion towards colleagues
- Acting in a responsible fashion towards society
- Behaving in a reliable and dependable way
- Showing altruism towards patients
- Being receptive to constructive criticism
- Having a positive attitude towards professional development
- Working well as a member of a team
- Reflecting on one's actions with a view to improvement
- Being attentive to the needs of colleagues
- Behaving with composure
- Being aware of one's limitations as a practitioner
- Providing advice to patients and colleagues when required
- Being adaptable to changes in the workplace
- Showing compassion towards one's patients
- Functioning according to the law
- Being punctual
- Treating other healthcare professionals fairly and without prejudice
- Treating other healthcare professionals fairly and without prejudice
- Being empathetic when caring for patients
- Being able to manage situations where there is a conflict of interest
- Treating colleagues fairly and without prejudice
- Not using one's professional status for personal gain
- Avoiding a cynical approach in one's job
- Being sensitive to the cultural background of colleagues and patients
- Making effective use of the resources available
- Being sound in judgment and in decision making
- Acting with confidence in one's duties
- Taking a dedicated approach to one's work
- Working with one's colleagues towards common goals
- Showing leadership skills and initiative
- Being accessible to patients
- Looking after one's own health and well-being
- Being accessible to colleagues
- Having the skills to train colleagues if required

Items with median rating of 3

- Conforming to social norms
- Having a good sense of humour
- Being physically fit
- Being well read outside the professional arena
- Being mindful of one's personal appearance

Items with median rating of 2

- Always being busy
- Earning a high salary

Items with median rating of 1

- Speaking with a refined accent
- Being physically attractive
- Owning a luxurious home
- Attending a prestigious school

4.3.4.3. Factor analysis

The sample adequacy for PCA, as measured by the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, was well above the acceptable value (minimum acceptable value is >0.5), and the non-correlations between items, as measured by Bartlett's Test of Sphericity, were sufficient and significant (<0.05) to perform a PCA (Dziuban & Shirkey 1974) (Table 31).

Table 31 – Sample adequacy and item suitability in the survey for a principal component analysis

Test		Values
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.977
Bartlett's Test of Sphericity	Approx. Chi-Square	24717.421
	df	1485.000
	Sig.	0.000

The principal component analysis generated six components and they explained more than 50% of the variance (Table 32). As a rule of thumb, the component structure deemed acceptable is > 50% of the variance (Starkweather & Herrington 2011).

Table 32 - Variance explained by the six principal components

factors	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	18.026	32.775	32.775	18.026	32.775	32.775	9.479	17.234	17.234
2	5.065	9.209	41.984	5.065	9.209	41.984	7.059	12.835	30.069
3	1.420	2.582	44.566	1.420	2.582	44.566	4.820	8.764	38.833
4	1.279	2.326	46.891	1.279	2.326	46.891	4.115	7.482	46.315
5	1.137	2.067	48.958	1.137	2.067	48.958	1.364	2.480	48.795
6	1.040	1.891	50.850	1.040	1.891	50.850	1.130	2.054	50.850

The majority of the items were clustered under four factors (Factors 1 -4) with the high factor co-efficient. By reviewing the items, factors one, two and four were named as *workmanship* (doctors' relationship with colleagues and other healthcare professionals), *clinicianship* (doctors' relationship with patient) and *citizenship* (doctors' behaviour in society) respectively. Factor three represented the non-important attributes of professionalism as perceived by the general public. Factors five and six were represented by only a single item each with an acceptable factor coefficient. Factors with fewer than three items and generally poor factor coefficients are regarded as trivial factors; the test developers can make sensible decisions on such factors and the items loaded on them (Brown 2009a). It is also advised that factor analysis should be used as guidance but not as a rule (Clark & Watson 1995). Therefore, the following changes were made: as 'being sensitive to the cultural background of colleagues and patients' was rated as 'very important' by the public it was grouped under *clinicianship*; and as 'being mindful of one's personal appearance' was not deemed to be important to the public and, therefore, omitted altogether. (Table 33)

Table 33 – Varimix rotated component matrix generated by principal component analysis of the public responses to 55 items and subscales (latent variables) identified

(Highest factor coefficient for each item is indicated in bold.)

	Component					
	1	2	3	4	5	6
Workmanship						
Respecting colleagues	**0.691**	0.325	0.084	0.125	-0.072	0.098
Treating colleagues fairly and without prejudice	**0.662**	0.283	0.071	0.221	-0.105	0.188
Being attentive to the needs of colleagues	**0.661**	0.250	0.228	0.127	0.011	0.114
Working well as a member of a team	**0.656**	0.226	0.065	0.187	0.156	0.005
Acting in a responsible fashion towards colleagues	**0.653**	0.237	0.129	0.240	-0.129	0.158
Treating other healthcare professionals fairly and without prejudice	**0.652**	0.278	0.042	0.207	-0.080	0.202
Being accessible to colleagues	**0.647**	0.299	0.183	0.066	0.082	-0.024
Working with one's colleagues towards common goals	**0.642**	0.319	0.068	0.133	0.068	0.083
Being adaptable to changes in the workplace	**0.610**	0.232	0.174	0.188	0.106	-0.006
Having the skills to train colleagues if required	**0.579**	0.177	0.104	0.070	0.315	-0.116
Being able to manage situations where there is a conflict of interest	**0.571**	0.333	0.104	0.235	-0.005	-0.145
Having a positive attitude towards professional development	**0.566**	0.281	0.172	0.224	0.105	0.036
Reflecting on one's actions with a view to improvement	**0.531**	0.316	0.165	0.246	0.117	0.044
Showing leadership skills and initiative	**0.522**	0.171	0.302	0.198	0.221	-0.113
Being receptive to constructive criticism	**0.515**	0.343	0.098	0.180	0.074	-0.078
Making effective use of the resources available	**0.509**	0.302	-0.057	0.253	0.104	-0.149
Being aware of one's limitations as a practitioner	**0.502**	0.176	0.021	0.283	-0.097	-0.233
Acting in a responsible fashion towards society	**0.454**	0.209	0.291	0.246	0.083	0.313
Acting with confidence in one's duties	**0.448**	0.301	0.046	0.419	0.058	-0.169
Looking after one's own health and well-being	**0.412**	0.239	0.236	0.239	0.215	-0.050
Not using one's professional status for personal gain	**0.308**	0.134	0.088	0.236	-0.367	-0.026
Being sensitive to the cultural background of colleagues and patients	***0.463***	0.257	0.104	0.138	0.061	**0.497**

Clinicianship						
Respecting patients autonomy	0.226	**0.703**	-0.042	0.237	-0.006	0.077
Being empathetic when caring for patients	0.271	**0.700**	-0.019	0.096	0.102	0.044
Showing compassion towards one's patients	0.305	**0.697**	0.003	0.166	0.048	0.031
Being attentive to the needs of patients	0.206	**0.683**	-0.014	0.273	-0.011	-0.012
Being accessible to patients	0.358	**0.643**	0.002	0.007	0.084	-0.168
Treating patients fairly and without prejudice	0.269	**0.590**	-0.149	0.253	-0.067	0.107
Acting in a responsible fashion towards patients	0.237	**0.582**	-0.098	0.340	-0.133	0.037
Providing advice to patients and colleagues when required	0.388	**0.558**	-0.012	0.121	0.112	-0.031
Behaving in a reliable and dependable way	0.394	**0.517**	0.002	0.332	-0.027	-0.167
Communicating in a clear and effective manner	0.370	**0.509**	0.007	0.309	-0.060	-0.108
Showing altruism towards patients	0.246	**0.504**	0.248	0.056	0.087	0.181
Respecting patient's confidentiality and privacy	0.171	**0.432**	-0.222	0.368	0.042	0.111
Avoiding a cynical approach in one's job	0.330	**0.425**	0.146	0.164	0.132	0.081
Behaving with composure	0.389	**0.413**	0.163	0.315	0.001	-0.012
Citizenship						
Adhering to professional rules and regulations	0.229	0.219	-0.057	**0.661**	-0.044	0.134
Functioning according to the law	0.218	0.230	-0.002	**0.652**	0.047	0.139
Avoiding substance or alcohol misuse	0.147	0.150	0.025	**0.502**	0.231	0.080
Behaving honestly and with integrity	0.188	0.427	-0.096	**0.497**	-0.086	-0.061
Being sound in judgment and in decision making	0.299	0.342	-0.015	**0.463**	-0.076	-0.211
Taking a dedicated approach to one's work	0.387	0.362	0.037	**0.437**	-0.082	-0.106
Being accountable for one's actions	0.369	0.365	-0.075	**0.428**	0.004	-0.126
Being punctual	0.337	0.290	0.224	**0.341**	0.215	-0.124
Non-important attributes of professionalism						
Being physically attractive	0.055	-0.068	**0.807**	-0.113	0.020	0.008
Owning a luxurious home	-0.013	-0.136	**0.784**	-0.076	-0.166	-0.028
Attending a prestigious school	0.049	-0.039	**0.753**	-0.064	-0.009	0.104
Speaking with a refined accent	0.031	0.011	**0.734**	0.025	0.018	0.048
Earning a high salary	0.138	-0.032	**0.627**	0.046	-0.061	-0.084
Always being busy	0.268	0.036	**0.572**	-0.042	0.288	-0.028
Being well read outside the professional arena	0.162	0.050	**0.558**	0.099	0.351	0.093
Being physically fit	0.298	0.076	**0.425**	0.118	0.355	-0.201
Conforming to social norms	0.272	0.023	**0.414**	0.122	0.393	0.186
Having a good sense of humour	0.358	0.172	**0.394**	-0.043	0.064	-0.326
Being mindful of one's personal appearance	0.278	0.125	***0.383***	0.346	**0.390**	0.013

No new attributes of professionalism were proposed by the general public by the way of free text comments. As indicated in the two questions with a Likert scale, the reaction of the public towards the survey was very positive and the majority enjoyed taking part in the survey (Table 34).

Table 34 - The evaluation of the survey by the public

Evaluation question	Response category					Total No. of respondents
Did you enjoy taking this survey? (1= not at all and 5= very much)	1	2	3	4	5	
	5 (0.5%)	14 (1.5%)	146 (15.3%)	237 (24.8%)	552 (57.9%)	954 (100%)
How likely are you to participate in the next survey? (1= very unlikely and 5= very likely)	1	2	3	4	5	
	3 (0.3%)	3 (0.3%)	97 (10.2%)	183 (19.2%)	668 (70%)	954 (100%)

Only 3.8% of the respondents included qualitative comments. The emerging themes indicated that it was an interesting survey in terms of format, length and content, but pointed out a few areas to be improved upon (Table 35).

Table 35 - Positive and negative comments made by the public about the survey

Positive aspects
Liked the drag and drop format *e.g. 'A different way of filling out a questionnaire, I really liked it, thank you.'*
The length is short *e.g. 'Easy survey and nice and short'*
Easy to complete *e.g. 'Liked this survey it was very easy.'*
Interesting *e.g. 'The questions asked were very interesting and informative especially in todays' society'*
Areas to be improved
Appear to have repetitive questions *e.g. 'A few of the questions were repeated'*
Should have had the option of going 'back' *e.g. 'Even though this was much easier to focus on, I don't like the method of picking answers. It's just too brief and if I think about it and change my mind afterwards, I don't have that option.*
No provision to indicate 'retired' under occupation *e.g. 'I was asked in what area I work. No provision for "retired" hence had to answer for the area in which I used to work.'*
Meaning of some words was not clear *e.g. 'Some of the questions struck me as having an ambiguous edge to them that could be left open to interpretation.'*
Format should have been different *e.g. 'The dragging requirement was too long and became annoying'*

4.3.5. Discussion

This is the first UK representative survey of the general public on medical professionalism in the published literature. Forty four out of the 46 evidence-based attributes of professionalism were identified as either 'very' or 'extremely' important. These items basically clustered under three components; the doctors' relationship with colleagues or other healthcare professionals (workmanship) and patients (clinicianship), and their behaviour in society (citizenship).

4.3.5.1. The methodology

The quota sampling techniques helped achieve a representative sample of the general public; the quotas adequately represented the major demographic groups, which are frequently considered important in a quota sample survey of the public (Smith 1983). The number of survey respondents was well above the threshold for subsequent statistical analyses intended in this project, e.g. a factor analysis requires more than 200 respondents to achieve a credible outcome (Clark & Watson 1995), and was comparatively high when compared to a random national telephone survey conducted among US patients to evaluate their perception on professionalism (954 in the present study versus 415 in the US survey) (Green *et al.* 2009). Accessing a convenience sample of the public through a database of volunteers, however, may have caused significant bias compared to a random sample (Cohen & Manion 1994, pp.86-89; Eun-ok & Wonshik 2011). This bias appear to be minimal with quota sampling compared to a simple convenience sampling, as the former addresses the issue of underrepresented subgroups in the study population, which is an important drawback associated with the latter (Eun-ok & Wonshik 2011). In addition, although the basic demographic characteristics (e.g. age, gender, social class) of the UK general public were represented adequately in the quota sample, the literacy level of the public who signed up to participate in research projects and surveys may be different to the general public as a whole. Therefore, the perceptions on professionalism among the two groups may be different. However, given that the primary purpose of this survey was conceptualising professionalism in the perspective of the general public in order to develop an assessment method, using a convenient representative sample rather

than a random sample could be further rationalised; sampling technique reflects and complies with the purpose of the study (Cohen & Manion 1994, p.87).

4.3.5.2. The findings

As the most frequent interaction between patients and doctors takes place during doctor-patient consultations, it is unsurprising and compatible with other research conducted using patients both in the UK (Miles & Leinster 2010) and outside (Green *et al.* 2009; Wagner *et al.* 2007; Wiggins *et al.* 2009) that attributes related to patient care were given a high priority by the public. However, unlike the above mentioned research, certain attributes relating to the interaction of doctors with co-workers and society also featured highly in this study. Green *et al* (2009) conducted a national survey of 415 patients in the US and also observed that the conceptualisation of professionalism by patients was not confined to the attributes of doctor-patient relationship.

Two literature-based items were rated as less important by the public in this survey. 'Conforming to social norms' may be an uncommonly thought of attribute of doctors' professionalism among the public. However, personal appearance of doctors (e.g. personal hygiene, dress code) emerges frequently as an attribute of professionalism in both quantitative (Green *et al.* 2009) and qualitative (Miles & Leinster 2010; Wagner *et al.* 2007; Wiggins *et al.* 2009) surveys of patients . There is evidence to suggest that patients equate dress code with personal hygiene, professional identity and scientific thoroughness (Gooden *et al.* 2001). Therefore, the lower importance given to personal appearance by the UK public in this survey was not compatible with most of the literature.

The results indicated that there are three facets to medical professionalism. However, Miles and Leister (2010) in an interview survey of 33 patients in the region of East Anglia, discovered that almost all professional attributes identified by their respondents were related to good communication and interpersonal skills (e.g. friendliness, listening patiently, caring attitude, treating patients with respect) and not related to team-working or upholding competence. Patients in the US do not appear to be different. According to Wagner *et al* (2007), who conducted focus groups with 11 patients, patients' conceptualisation of professionalism is largely related to doctor-patient relationship. These discrepancies may well be attributable to the difference between the study designs. Miles & Leister (2010) and Wagner *et al* (2007) analysed spontaneous qualitative comments by patients, which depended on the broadness of each patient's understanding of professionalism. In the present study, the researcher analysed public responses to a given set of attributes, which may introduce or broaden the understanding of members of the public about the topic. The findings of Green *et al* (2009), who employed a similar approach to the one used in this study observed a broad conceptualisation of professionalism among patients. Alternatively, this may support the notion that patients have different views from the general public. Both these suggestions may be explained by the fact that the perceptions of lay people can be influenced by recent experiences or exposure to salient information (Geer 1991).

The three-facet conceptualisation of professionalism, however, is comparable with similar conceptualisations generated through the thematic analysis of literature, cumulative

response of patients, students, postgraduate trainees, and doctors. For example, the domains of inter-personal (relationship with patients and co-workers, e.g. altruism, honesty, compassion, reliability, communication skills), intra-personal (personal characteristics for effectiveness and efficiency, e.g. lifelong learning, morality, non-cynical approach, self-care) and pubic professionalism (maintaining public trust, e.g. accountability, commitment to ethical /moral codes, keeping up-to-date, maintaining standards), generated by van de Camp *et al* (2004) based on a thematic analysis of literature followed by expert validation, is somewhat compatible with this model. Although the relationship with patients and co-workers cluster together under the interpersonal domain in the van de Camp *et al* model, it features as two distinct groups in the three-facet model. Similarly, the intra-personal domain of the van de Camp *et al* model appears to be represented by both clinicianship and workmanship facets. The public domain of the van de Camp model, however, is largely compatible with the citizenship facet. A qualitative analysis of written comments of supervising clinicians on the professionalism of Michigan medical students has produced initiative, relationship with the team, patient skills, work habits, self-improvement and composure as emerging themes (Frohna & Stern 2005). It appears that except for patient skills, which overlap with clinicianship, all other themes represent the workmanship facet of the current study. No theme appears to represent the citizenship facet. An analysis of narratives written by Indiana medical students on incidents related to their learning of professionalism helped Karnieli-Miller *et al* (2010) to identify the following themes; manifesting respect/disrespect in clinical encounters with patients, families, co-workers and

colleagues.; managing communication challenges with patients; demonstrating responsibility, pride, knowledge and thoroughness; spending time taking care of patients, patients' education and understanding; altruism; communicating and working in teams. Again, these themes overlap considerably with the workmanship and the clinicianship facets but not with the citizenship facet. Similarly, a cumulative thematic analysis of focus group discussions with doctors, students and patients in the region of East Anglia, UK (Miles & Leinster 2010), has generated three themes; good team-working, upholding competence and good communication and interpersonal skills which are compatible with workmanship and clinicianship facets, but not with the citizenship facet. Behaviour of doctors in society, however, is an important aspect of professionalism (GMC 2009) though it has not been a commonly emerging theme in the research literature. Therefore, the three-facet model is constituted of clinicianship and workmanship, which are recurring themes, and citizenship, which is quite novel to the professionalism literature, but deemed to be important in conceptualising modern-day professionalism.

4.3.5.3. Conclusions

This part of the study reliably gathered the public perception on medical professionalism and generated a sound conceptual model. This model is broader than the existing ones in the literature and appears to encompass all the important aspects of professionalism. The three facets of the model (clinicianship, workmanship and citizenship), therefore, would be taken forward to be used as the basic template for assessing professionalism culture.

4.3.5.4. Publications and presentations related to the public survey

- Chandratilake MN, McAleer J, Gibson J and Roff S. (2010) Medical Professionalism: what does public think? *Clinical Medicine*. 10; 1- 6. (Appendix IV)
- Chandratilake M, McAleer S, Gibson J & Roff S. Clinicianship, workmanship, citizenship: the public's model of medical professionalism. Presented at the Scottish School of Primary Care International Annual Conference, 27 -28 April 2010, Creiff Hydro Hotel, Creiff, UK
- Chandratilake M, McAleer S, Gibson J & Roff S. Viewing medical professionalism from the perspective of the general public. Presented at the Annual Conference of the Association for Medical Educators in Europe (AMEE), 5 – 8 September 2010, Scottish Exhibition and Conference Centre, Glasgow, UK.
- Chandratilake M, Patient-centred professionalism. Presented at the Annual Meeting of Academy of Medical Educators, 26 – 27 January 2011, Grosvenor Square Marriott Hotel, London, UK.

4.4. The professionals' survey

This survey was conducted specifically to identify the most important attributes of professionalism to medical professionals in the UK. Unlike the general public, medical professionals can be considered as 'experts' in a process of developing assessments / psychological measures (De Villiers *et al.* 2005; Polit & Beck 2006). Therefore, the Content Validity Index (CVI) was used to determine the items of most importance to the medical professionals in the UK. In this empirical measure of content validity, surveyed responses of an expert group are analysed to determine the essentialness of each item to represent an underlying construct (Johnston & Wilkinson 2009; Polit & Beck 2006), medical professionalism in this instance. CVI has been recommended in the development of assessments (Johnston & Wilkinson 2009) and psychological measures (Polit & Beck 2006; Rubio *et al.* 2003).

4.4.1. Sampling and recruitment

'Experts' in the context of establishing CVI can be either people, who are well-experienced in the topic by working or publishing, or are representatives of stakeholder groups, to whom the topic is most relevant (Rubio *et al.* 2003). Although five to 20 experts are recommended for practicality reasons (Lynn 1986), the larger the number the more credible the result (Rubio *et al.* 2003). Therefore, it was decided to invite the medical professionals enrolled in the Postgraduate Certificate, Diploma and Masters Courses at the Centre for Medical Education, University of Dundee, to participate. As professionalism may be culture-sensitive all 2183 international enrolees were invited to participate in the

survey with the intention of contributing to the professionalism literature by examining geographical and cultural similarities and differences. However, as the current measure of institutional professionalism culture is focused on the UK undergraduate setting, only the response of UK medical practitioners were considered for the purpose of this project. A paper on the cultural and geographical similarities and differences identified in the survey, but excluded from this thesis, has already been published (Appendix V).

The inclusion of a selected group of medical professionals involved in teaching and/or clinical work as the panel of experts was appropriate. In the development of the Jefferson scale of empathy a similar group of 100 experts was used (Hojat *et al.* 2001). There can be variations in the level of knowledge and experience in professionalism within this group. These may affect the responses and the final outcome, especially with small number of panellists (Lynn 1986). However, with a large panel the potential bias caused by the varying levels of expertise in a consensus process, could be overlooked (De Villiers *et al.* 2005).

4.4.2. Data collection

The survey content and rating scale were the same as for the public survey; the respondents were expected to rate the 55 attributes on a five-point Likert scale indicating the importance of each item as a professional attribute for doctors (5= extremely important, 4 = very important, 3 = somewhat important, 2 = slightly important, 1= unimportant). The question asked of the professionals was: how important are the following attributes of medical professionalism? In addition to the items, certain

demographic information was requested at the end of the survey (gender, age group, professional involvement, teaching involvement, country of residence).

Unlike the public survey, the medical professionals were surveyed using both paper-based and online formats of the inventory. Although the order of items was the same in the public and professionals' survey, the format used in both versions of the professionals' survey contained multiple items on the same paper/screen page with one option out of the five to be ticked. The paper-based version, with a self-addressed envelope, was posted, and the online version (developed using Bristol Online Surveys), as a web-link, was emailed. The potential respondents were specifically advised to respond to only one version of the survey; either postal or email. No incentives were offered for participating in the survey. One hundred and twenty six persons were inaccessible as both electronic and postal mails sent to them were undeliverable, thus reducing the number of potential respondents to 2057. The submission of responses in both versions was anonymous. An email reminder was sent one week after the launch. The results were analysed 11 weeks after launching the survey. The rationale for using this strategy in the professionals' survey is discussed below.

a. Incentives

In a survey of health professionals, around a 33% response rate can be anticipated without providing incentives (Deehan *et al.* 1997). Substantial monetary incentives may enhance this rate at least moderately (Deehan *et al.* 1997; Delnevo *et al.* 2004). As the cohort selected was large (more than 2000 potential participants), the anticipated number

of responses without incentives (around 600 if the response rate is 30%) would be more than adequate for a validation study (Rubio *et al.* 2003). In the development of the Jefferson scale of empathy, without incentives, 55 out of 100 selected experts in the US responded (Hojat *et al.* 2001). Therefore, after considering the literature evidence and with the limited funds available for the project, no incentives for participation were offered to the medical professionals.

b. Reminders

The Ethics Committee advised the use of only one reminder. The literature evidence also suggested that only the first reminder increases the response rate significantly (Cook *et al.* 2009). Therefore, a single reminder was sent to all participants by email one week after the launch, which may be the most effective time to send a reminder (Deutskens *et al.* 2004).

Six hundred and sixty enrolees responded to the survey (overall response rate = 32%). Seventy six of the respondents were nurses, dentists and other healthcare professional, and 584 medical practitioners. The responses of 368 out of 1417 of UK medical practitioners enrolled in the course were used for the purpose of this validation study (response rate among UK medical practitioners = 25.9%). (Table 36)

Table 36 - Geographical origin of the medical professionals, who participated in the validation survey

	United Kingdom	Other European country	Asian country	Australasian county	North American country	South American country	African country	Total
Frequency	368	56	70	20	52	6	12	584
%	63	9.6	12	3.4	8.9	1	2.1	100

4.4.3. Data analysis

In this study, the Content Validity Index (CVI) (Lynn 1986) supplemented with multi-rater *kappa* statistics was used in analysing the responses to the survey.

The CVI of an item is the proportion of participating group members, who indicate that a particular item is 'essential' out of the total number in the group (Polit & Beck 2006). For example, if eight out of 10 members agree that a particular item is essential the CVI is 0.8. What points of the rating scale should be included in the 'essential' category is a decision to be made by the measure developers (Johnston & Wilkinson 2009; Lynn 1986). In this study, 'extremely' and 'very' important were classified as 'essential'. Given that the number of raters in the reference group are more than five, the general agreement is that the items which indicate CVIs of ≥ 0.78 (Lynn 1986), i.e. items, which are rated by 78% or more of the respondents as either 'very' or 'extremely' important, can be considered as essential.

As CVI is a proportion, the agreement calculated in terms of the CVI can be due to chance. Calculation of *kappa* values helps exclude the chance element from overall agreement among multiple raters (Kianifard 1994; Polit & Beck 2006). *Kappa* can range from -1.0 to

+1.0, with -1.0 indicating perfect disagreement below chance, 0.0 indicating agreement equal to chance, and 1.0 indicating perfect agreement above chance (Randolph 2008). Within 95% confidence interval, a rule of thumb is that a *kappa* of 0.70 or above indicates adequate inter-rater agreement (Randolph 2008). The *kappa* value for the set of 'essential' items, which were determined based on the CVI values, was subsequently calculated using the online multi-rater *kappa* calculator (Randolph 2008). The same response categories and number of respondents from each geographical region were used for this purpose. Free-margin multi-rater kappa is preferred to fixed-margin multi-rater kappa when the raters were not forced to choose a fixed number of items to each category (Randolph 2005). The CVI or CVI supplemented with *kappa* values was used to establish content validity of several healthcare assessments and measurements (Polit & Beck 2006).

4.4.4. Results

The results of the professionals' survey are presented as follows:

- Demographic characteristics of the sample
- Comparison of public and professionals' responses
- Identifying items for the final measure

4.4.4.1. Demographic characteristics of the sample

The sample of 368 UK respondents consisted of almost equal number of males and females, and undergraduate and postgraduate teachers. The majority of the respondents were aged between 30 and 50. (Table 37)

Table 37 - Demographic characteristics of UK medical professionals, who responded to the validation survey

Demographic category	Demographic group	Frequency	%
Gender	Male	167	45.4
	Female	198	53.8
	Prefer not to say	3	0.8
Age group	< 30 years	63	17.1
	30 - 50 years	254	69.0
	> 50 years	49	13.3
	Prefer not to say	2	0.5
Professional engagement	Mainly clinical	345	93.8
	Mainly non-clinical	23	6.3
Teaching involvement	Mainly undergraduate	168	45.7
	Mainly postgraduate	177	48.1
	No teaching involvement	23	6.3

4.4.4.2. Comparison of public and professionals' responses

The median responses indicated that there were no significant differences between the pubic and professionals' responses in determining the essentialness of attributes However, 'avoiding cynical approach to one's job' was rated as moderately important by the public but as 'very important' by the professionals (Table 38).

Table 38 - Comparison of importance placed by the public and professionals on the 55 items

Items with median 4 or 5 in both the public and professionals' surveys

- Respecting patients autonomy
- Being accountable for one's actions
- Behaving honestly and with integrity
- Respecting patient's confidentiality and privacy
- Communicating in a clear and effective manner
- Acting in a responsible fashion towards patients
- Behaving in a reliable and dependable way
- Being aware of one's limitations as a practitioner
- Showing compassion towards one's patients
- Treating patients fairly and without prejudice
- Avoiding substance or alcohol misuse
- Respecting colleagues
- Acting in a responsible fashion towards colleagues
- Acting in a responsible fashion towards society
- Showing altruism towards patients
- Being receptive to constructive criticism
- Having a positive attitude towards professional development
- Adhering to professional rules and regulations
- Working well as a member of a team
- Reflecting on one's actions with a view to improvement
- Being attentive to the needs of colleagues
- Behaving with composure
- Being attentive to the needs of patients
- Providing advice to patients and colleagues when required
- Being adaptable to changes in the workplace
- Functioning according to the law
- Being punctual
- Treating other healthcare professionals fairly and without prejudice
- Being empathetic when caring for patients
- Being able to manage situations where there is a conflict of interest
- Treating colleagues fairly and without prejudice
- Not using one's professional status for personal gain
- Being sensitive to the cultural background of colleagues and patients
- Making effective use of the resources available
- Being sound in judgment and in decision making
- Acting with confidence in one's duties
- Taking a dedicated approach to one's work
- Working with one's colleagues towards common goals
- Showing leadership skills and initiative
- Being accessible to patients
- Looking after one's own health and well-being
- Being accessible to colleagues
- Having the skills to train colleagues if required

Items with median 3 in both the public and professionals' surveys

- Conforming to social norms
- Having a good sense of humour
- Being physically fit
- Being well-read outside one's professional area
- Being mindful of one's personal appearance

Items with median 1 or 2 in both the public and professionals' surveys

- Earning a high salary
- Speaking with a refined accent
- Being physically attractive
- Always being busy
- Owning a luxurious home
- Having attended a prestigious school before going to university

Item which demonstrated differences in medians with practical significance

- Avoiding a cynical approach in one's job (The professionals 4 and the public 3)

4.4.4.3. Identifying items for the final measure

The examination of agreement among raters on the essentialness of items as attributes of professionalism revealed that thirty evidence-based items met the necessary CVI value (≥ 0.78) (Table 39). *Kappa* for 30 essential items was 0.77, which indicated that the agreement on those items among the raters was not due to chance.

Table 39 – The items with acceptable agreement (CVI ≥ 0.78) among UK medical professionals on the essentialness of professional attributes

	No. of 'extremely important' responses (nR5)	No. of 'very important' responses (nR4)	No. of 'essential' responses (nR5+nR4)	Content Validity Index (CVI) (nR5+nR4)/total number of respondents
Respecting patients' autonomy	293	72	365	0.99
Being accountable for one's actions	272	94	366	0.99
Behaving honestly and with integrity	320	46	366	0.99
Acting in a responsible fashion towards patients	253	111	364	0.99
Being aware of one's limitations as a practitioner	245	118	363	0.99
Respecting patient's confidentiality and privacy	271	90	361	0.98
Treating patients fairly and without prejudice	249	113	362	0.98
Respecting colleagues of the same profession	183	169	352	0.96
Communicating in a clear and effective manner	214	140	354	0.96
Being sound in judgment and in decision making	170	182	352	0.96
Acting in a responsible fashion towards colleagues	180	171	351	0.95
Behaving in a reliable and dependable way	196	149	345	0.94
Being attentive to the needs of patients	166	179	345	0.94
Treating other healthcare professionals fairly and without prejudice	156	190	346	0.94
Functioning according to the law	194	147	341	0.93
Treating colleagues fairly and without prejudice	154	189	343	0.93
Reflecting on one's actions with a view to improvement	142	193	335	0.91
Showing compassion towards one's patients	200	135	335	0.91
Being empathetic when caring for patients	165	169	334	0.91
Working well as a member of a team	146	182	328	0.89
Having a positive attitude towards professional development	124	198	322	0.88
Being able to manage situations where there is a conflict of interest	83	239	322	0.88
Providing advice to patients and colleagues when required	129	190	319	0.87
Being receptive to constructive criticism	105	212	317	0.86
Avoiding substance or alcohol misuse	193	123	316	0.86
Not using one's professional status for personal gain	175	134	309	0.84
Taking a dedicated approach to one's work	98	200	298	0.81
Adhering to professional rules and regulations	112	183	295	0.8
Making effective use of the resources available	95	196	291	0.79
Acting in a responsible fashion towards society	116	172	288	0.78

The remaining 25 demonstrated below-acceptable agreement (Table 40).

Table 40 - The items with below acceptable agreement (CVI < 0.78) among UK medical professionals on the essentialness of professional attributes

	No. of 'extremely important' responses (nR5)	No. of 'very important' responses (nR4)	No. of participant who indicated that the item is 'essential' (nR5+nR4)	Content Validity Index (CVI) (nR5+nR4)/total number of respondents
Working with one's colleagues towards common goals	81	199	280	0.76
Being accessible to patients	90	189	279	0.76
Having the skills to train colleagues if required	99	179	278	0.76
Being punctual	63	213	276	0.75
Being sensitive to the cultural background of colleagues and patients	100	177	277	0.75
Showing leadership skills and initiative	79	197	276	0.75
Acting with confidence in one's duties	73	201	274	0.74
Looking after one's own health and well-being	107	167	274	0.74
Being attentive to the needs of colleagues	64	195	259	0.70
Being accessible to colleagues	68	180	248	0.67
Showing altruism towards patients	80	162	242	0.66
Being adaptable to changes in the workplace	63	179	242	0.66
Behaving with composure	36	175	211	0.57
Avoiding a cynical approach in one's job	54	120	174	0.47
Being mindful of one's personal appearance	18	110	128	0.35
Having a good sense of humour	18	72	90	0.24
Being physically fit	9	68	77	0.21
Conforming to social norms	12	56	68	0.18
Being well-read outside one's professional area	15	48	63	0.17
Earning a high salary	10	30	40	0.11
Speaking with a refined accent	7	22	29	0.08
Always being busy	4	17	21	0.06
Being physically attractive	4	8	12	0.03
Owning a luxurious home	1	8	9	0.02
Having attended a prestigious school before going to university	2	4	6	0.02

Overall percentage agreement on the 30 essential attributes of professionalism among the 368 medical professionals was 84% and was not due to chance (*kappa* 0.77).

All thirty 'essential' attributes also appeared in the list of 'most important' attributes as rated by the UK public. Therefore, those items were eligible to be included in the final measure. However, after reviewing the items, it was decided to exclude 'acting in a responsible fashion towards colleagues' and 'acting in a responsible fashion towards society' as they appeared to be broad and less specific compared to the other items. Broad and less specific items may affect the psychometric rigour of a measure and making such subjective judgements to improve the relevance, specificity and quality of items is warranted in developing measures of this kind (DeVellis 2003). Twenty eight items were, therefore, included to represent the three facets of professionalism identified in the public survey. (Table 41)

Table 41 - Items identified for the first field test of the measure by survey distillation

Items essential to the public	Items essential to professionals	Included in the final measure
1. Respecting patients autonomy	√	√
2. Behaving honestly and with integrity	√	√
3. Respecting patient's confidentiality and privacy	√	√
4. Acting in a responsible fashion towards patients	√	√
5. Adhering to professional rules and regulations	√	√
6. Being attentive to the needs of patients	√	√
7. Treating patients fairly and without prejudice	√	√
8. Avoiding substance or alcohol misuse	√	√
9. Being accountable for one's actions	√	√
10. Respecting colleagues of the same profession	√	√
11. Communicating in a clear and effective manner	√	√
12. Behaving in a reliable and dependable way	√	√
13. Being receptive to constructive criticism	√	√
14. Having a positive attitude towards professional development	√	√
15. Working well as a member of a team	√	√
16. Reflecting on one's actions with a view to improvement	√	√
17. Being aware of one's limitations as a practitioner	√	√
18. Providing advice to patients and colleagues when required	√	√
19. Showing compassion towards one's patients	√	√
20. Functioning according to the law	√	√
21. Treating other healthcare professionals fairly and without prejudice	√	√
22. Being empathetic when caring for patients	√	√
23. Being able to manage situations where there is a conflict of interest	√	√
24. Treating colleagues fairly and without prejudice	√	√
25. Not using one's professional status for personal gain	√	√
26. Making effective use of the resources available	√	√
27. Being sound in judgment and in decision making	√	√
28. Taking a dedicated approach to one's work	√	√
29. Acting in a responsible fashion towards colleagues	√	
30. Acting in a responsible fashion towards society	√	
31. Working with one's colleagues towards common goals		
32. Being adaptable to changes in the workplace		
33. Being attentive to the needs of colleagues		
34. Behaving with composure		
35. Showing altruism towards patients		
36. Showing leadership skills and initiative		
37. Acting with confidence in one's duties		
38. Being punctual		
39. Avoiding a cynical approach in one's job		
40. Being sensitive to the cultural background of colleagues and patients		
41. Being accessible to patients		
42. Looking after one's own health and well-being		
43. Being accessible to colleagues		
44. Having the skills to train colleagues if required		

Accordingly, 11 out of 21 items of workmanship, 10 out 14 items of clinicianship and seven out of nine items of citizenship were selected for the final measure (Table 42).

Table 42 - The final list of items considered under each facet of professionalism

Workmanship	Clinicianship	Citizenship
1. Respecting colleagues of the same profession	1. Respecting patients' autonomy	1. Being accountable for one's actions
2. Reflecting on your actions with a view to self-improvement	2. Behaving in a reliable and dependable way	2. Behaving honestly and with integrity
3. Having a positive attitude towards professional development	3. Communicating with patients in a clear and effective manner	3. Adhering to professional rules and regulations
4. Working well as a member of a team	4. Acting in a responsible fashion towards patients	4. Functioning according to the law
5. Being aware of own limitations	5. Being attentive to the needs of patients	5. Avoiding substance or alcohol misuse
6. Being receptive to constructive criticism	6. Providing advice to patients when required	6. Being sound in judgment and in decision making
7. Treating other healthcare professionals fairly and without prejudice	7. Showing compassion towards patients	7. Taking a dedicated approach to work
8. Being able to manage situations where there is a conflict of interest	8. Respecting patients' confidentiality and privacy	
9. Treating colleagues of the same profession fairly and without prejudice	9. Treating patients fairly and without prejudice	
10. Not using professional status for personal gain	10. Being empathetic when caring for patients	
11. Making effective use of the resources available		

4.4.5. Discussion

The attributes of professionalism, deemed to be essential to medical professionals in the UK, were identified by analysing survey responses of a convenience sample of 368 experts. The essentialness of attributes was determined by using the CVI. Based on the feedback from the public and professionals, 28 items were selected to represent the three facets of professionalism in the measure of professional culture.

4.4.5.1. The methodology used

As in the present study, convenience samples of medical professionals have been used historically as 'experts' in validating the measures of the professionalism environment. The experts, however, were not necessarily involved directly in teaching or researching the area concerned (Hojat *et al.* 2001; Roff *et al.* 1997). The use of a convenience sample of UK medical practitioners as 'experts' in this validation study, therefore, complies with these previous studies. However, the responses of a random sample of clincians in the UK, e.g. clinicians working in general practice or hospital setting, may be different to this selected group. The objective of this validity survey, however, was to identify the items to represent a conceptual framework of professionalism and not surveying doctors' perception of professionalism. Therefore, the use of the ratings of a selected group would be acceptable.

Although the number of respondents is more than adequate for the purpose of this study, the response rate may be below the average for surveys of healthcare professionals in general (Deehan *et al.* 1997).

The CVI supplemented with *kappa* statistics has been used in identifying items for many healthcare-related measures, e.g. osteopoporsis risk assessment tool (Wynd *et al.* 2003), and a tool for measuring the effectiveness of health education (Black *et al.* 2011). However, the CVI has not been used to select items and achieve the agreement of experts in any of the measures of professionalism culture considered in the literature review. With the intention of developing a professionalism assessment, a similar approach was used by Green *et al* (2009) to analyse survey results of physicians. In a validation survey of 214 US physicians, instead of 78% (the cut-off agreement as advised in CVI) Green *et al* (2009) arbitrarily used 75% as the acceptable agreement and made no attempts to exclude the possibility of agreement among raters by chance. Although given the high number of participants, selecting 75% may be permissible (Lynn 1986), not investigating the chance element in the observed agreement may have led to inflated percent agreement in the study by Green *et al* (2009) (Watkins & Pacheco 2000).

The use of a cut-off point informed by the literature, however, may have left out certain items which deemed to be important. Although this complies with the whole objective of selecting a practicable number of items with the highest consensus, numbers may not always reflect the suitability and relevance of items. However, in this instance, the items filtered by the process appeared to be well-representative of the underlying facets even in a subjective judgement.

As discussed above under the justification, the consensus survey method was the most appropriate method of determining the content validity. However, unlike in Delphi

technique there was no iterative dialogue with the reference group in the process of prioritisation of items, which may be a limitation of using the survey method.

4.4.5.2. The findings

In the present study, the importance placed on the 55 attributes of professionalism by the professionals was closely similar to the general public. These findings, therefore, are dissimilar to the study by Miles and Leinster (2010), where UK patients confined the scope of professionalism to doctor-patient relationship, but doctors expanded it to encompass clinical competence and teamwork. A questionnaire survey conducted in the US by Green *et al* (2009) revealed that patients put more importance than doctors on communicating with other healthcare professionals, exploring patients' needs and concerns, respecting colleagues, advising patients and colleagues, effective use of resources and cultural sensitivity.

Nevertheless, the 28 attributes of professionalism were considerably similar to the findings of several studies that examined the perspective of doctors on professionalism. For example, the top ten attributes of professionalism (upholding competence, respect, empathy, honesty and integrity, responsibility, collegiality, confidentiality, up-to-date knowledge, courteous, good communication) identified by the Canadian postgraduate trainees were almost identical to and representative of the 30 attributes, though the items in the former are broader than the items in the latter (Brownell & Cote 2001). The perspective of UK psychiatrists on professionalism was largely similar to the areas represented by the 30 items except for self-governance, an aspect of which the

psychiatrists felt to be important (Bhugra 2008). In a similar survey conducted in the UK, trainees in internal medicine and surgery emphasised the autonomy of clinicians as an areas of professionalism in addition to the areas related to patients, co-workers, professional bodies and society (Chard *et al.* 2006). Self-governance, however, is no longer important in the neo-professionalism (Wass 2006). The final list of professional attributes is largely similar to the attributes identified by the UK medical students (Monrouxe *et al.* 2011).

4.4.5.3. Conclusions

The public survey, the survey of professionals and the subsequent comparative analysis helped identify the 28 items, which were reasonably compatible with similar studies. These items represent the three facets of professionalism in the measure of institutional professionalism culture to be developed.

4.4.5.4. Publications and presentations

- **Chandratilake M**, McAleer S, Gibson J. (2012) Cultural similarities and differences in medical professionalism: a multi-region study. *Medical Education*, 46: 257–66. (Appendix V)

- **Chandratilake M**, McAleer S, Gibson J & Roff S. Personal, professional and cultural differences in the perception of healthcare professionalism. Presented at the Annual Meeting of the Association for the Study of Medical Education (ASME), 21 – 23 July 2010, Robinson College, Cambridge, UK.

Section C: Designing the measure of professionalism culture and determining its utility

In this section the research question 2, 'how can institutional professionalism of medical schools in the UK be measured quantitatively with appropriate validity, reliability and acceptability?' is addressed.

As discussed in the literature review, the culture of an institution manifests as the cumulative behaviour of its inhabitants (Hudelson 2004). The prevalence of the behaviours, (identified in the previous section) within a given institution, therefore, should provide a measure of its professionalism culture. Direct observations and surveys would be appropriate for this purpose (Scott *et al.* 2003a). Although both quantitative and qualitative data can be collected using these methods (Scott *et al.* 2003a), given the aim of the doctoral project, the focus is on gathering quantitative data.

The authenticity of observational data is higher than self-perceptions (Cohen & Manion 1994, pp.106-125). However, the observations should be implicit or secretive as the presence of observers may influence the behaviours under observation (Cohen & Manion 1994, pp.106-125). In addition to the ethical issues (Moore & Savage 2002), conducting an institution-wide observational study of behaviours to determine the professionalism culture would be overwhelmingly resource-intensive and not very practical. These difficulties become critical if repeated or frequent measures are to be planned. A practical alternative is the collation of the perceptions of multiple sources (e.g. staff and students of a medical school) on the frequency and intensity of professional/unprofessional

behaviours, and questionnaire surveys are commonly used for this purpose (Baldwin & Daugherty 2006, p.105). The perceived frequency or intensity of a given behaviour, however, may be affected by many factors. The questions targeting one's own behaviour may be affected adversely by the unreliability of self-reporting, the variability of reflective ability among people and the propensity to select socially desirable responses (Cohen & Manion 1994, pp.106-125 and pp. 83-104). Focusing questions on the institution or environment rather than self, however, tends to yield more credible responses (Baldwin & Daugherty 2006, p.105). Even if the questions are focused on the behaviours of other people in the institution, the frequency of a particular behaviour may be affected by a single event with high intensity and *vice versa*. Therefore, a theory-based system of gathering responses needs to be considered to enhance the credibility of responses (DeVellis 2003, pp.6-8).

The Theory of Planned Behaviour (TPB) (Ajzen 1991) was identified in the literature review (Section 2.5.2, p.1) as the suitable theoretical basis to develop a system of gathering data on the institutional professionalism culture. According to the TPB (Ajzen 1991), a voluntary behaviour takes place when there is an intention. The intention, in turn, is determined by personal attitude and what one perceives as the expectation of people around oneself towards the behaviour concerned, and the perceived capability (e.g. possessing the necessary knowledge and skills, availability of resources) of executing the behaviour (Figure 5, p.1). Therefore, by measuring attitudes, perceived expectations and perceived achievability in relation to a given behaviour, it is possible to predict its

prevalence more credibly and accurately than by gathering perceived frequency or intensity.

The next two chapters discuss the process of designing a measure with the attributes of professionalism identified in Section B using TPB as the basis (Chapter 5) and establishing its psychometric properties and acceptability by field-testing (Chapter 6).

Chapter 5 – Designing the measure of professionalism culture

In this chapter, the development of the measure of institutional professionalism culture using TPB is discussed.

5.1. Developing the response-gathering system based on TPB

According to the public survey findings, in determining professionalism, there are three key behavioural aspects to focus on; taking a patient-centered approach in doctor-patient encounters (clinicianship), showing team orientation in the workplace (workmanship), and behaving with a socially responsible manner in society (citizenship). Therefore, in the context of this study, the attitude, perceived social norm and perceived behavioural control towards these three behaviours should be measured. These key behaviours, however, are latent, and are difficult to measure directly (DeVellis 2003, pp.10). The items clustered under each of these key behaviours are directly observable manifestations of the respective latent behaviour and, cumulatively explain the prevalence of the latent behaviour, i.e. indexical items (DeVellis 2003, pp.10). For example, the 10 items, which represent the clinicianship, can be considered as the measurable manifestations of 'taking patient-centered approach in doctor-patient encounters'. The attitude, perceived social norm and perceived behavioural control towards each of the 10 items synthesises the attitude, perceived social norm and perceived behavioural control towards 'taking patient-centered approach in doctor-patient encounters'.

5.1.1. Describing behaviours under TPB

In a measure of behaviours underpinning TPB, the stem of the question contains the targeted behaviour and may be defined precisely in terms of Target, Action, Context, and Time (TACT) (Ajzen 2002). If TACT elements were to be incorporated to a behaviour identified in this study, it would be as follows:

Thinking about the current clinical rotation you are in (Time), respecting (Action) colleagues of the same profession (Target) in the workplace (Context) is……

However, in the context of this study, 'time' element is common to all 28 behaviours and 'context' element is common to the set of items, which represent a particular facet. In such an event, one or more TACT elements may be made generalised to multiple behaviours (Ajzen 2002). Therefore, *time* was generalised to all items and not included in the questionnaire but mentioned in the email which contained the link to the measure, and *context* was generalised to each facet. On the other hand, when behaviours imply a general rather than a specific context, such behaviours do not need a specified 'target' (Ajzen 2002). Accordingly, two behaviours (behaving honestly and with integrity and behaving in a reliable and dependable way) which were generalisable were not specified with targets. (Table 43)

Table 43 - Analysis of behaviours included in the professionalism culture measured against the TPB elements

Behaviours	Action	Target	Context
Respecting colleagues of the same profession	Respecting	colleagues of the same profession	In workplace
Reflecting on your actions with a view to self-improvement	Reflecting your actions	self-improvement	
Having a positive attitude towards professional development	Having a positive attitude	professional development	
Working well as a member of a team	Working well	member of a team	
Being aware of own limitations	Being aware	own limitations	
Being receptive to constructive criticism	Being receptive	constructive criticism	
Treating other healthcare professionals fairly and without prejudice	Treating fairly and without prejudice	other healthcare professionals	
Being able to manage situations where there is a conflict of interest	Being able to manage situations	conflict of interest	
Treating colleagues of the same profession fairly and without prejudice	Treating fairly and without prejudice	colleagues of the same profession	
Not using professional status for personal gain	Not using professional status	personal gain	
Making effective use of the resources available	Making effective use	resources	
Respecting patients' autonomy	Respecting autonomy	patients	During patient encounters
Behaving in a reliable and dependable way	Behaving reliable and dependable way	(with others – as implied)	
Communicating with patients in a clear and effective manner	Clear and effective communication	patients	
Acting in a responsible fashion towards patients	Acting in a responsible fashion	towards patients	
Being attentive to the needs of patients	Being attentive to needs	patients	
Providing advice to patients when required	Providing advice when required	patients	
Showing compassion towards patients	Showing compassion	towards patients	
Respecting patients' confidentiality and privacy	Respecting confidentiality and privacy	patients	
Treating patients fairly and without prejudice	Treating fairly and without prejudice	patients	
Being empathetic when caring for patients	Being empathetic	patients	
Being accountable for one's actions	Being accountable	one's actions	As a part of wider society
Behaving honestly and with integrity	Behaving honestly and with integrity	(with others – as implied)	
Adhering to professional rules and regulations	Adhering	professional rules and regulations	
Functioning according to the law	Functioning	the law	
Avoiding substance or alcohol misuse	Avoiding	substance or alcohol	
Being sound in judgment and in decision making	Being sound	judgment and decision making	
Taking a dedicated approach to work	Taking a dedicated approach	work	

5.1.2. Developing the response format based on TPB

All three components of TPB can be measured directly (e.g. asking participants about the attitude of the respondent towards the targeted behaviour itself) or indirectly (e.g. synthesising the attitude of the respondent towards the targeted behaviour through obtaining their perceptions related to the targeted behaviour, e.g. benefits of demonstrating the targeted behaviour) (Francis *et al.* 2004). Although researchers have measured these components indirectly for the purpose of developing TPB questionnaires (Francis *et al.* 2004), the direct measurement is preferred by the TPB theorist (Ajzen 2002). Therefore, it was decided to use 'direct measures' against the three TPB components in relation to each of the targeted behaviours.

Semantic differential scales, i.e. multiple scales with bipolar adjectives (e.g. good – bad, valuable-worthless), are commonly advocated for the direct measurement of cognitive constructs such as attitudes (Ajzen 2002; DeVellis 2003, p.80), and have been used in many healthcare-related questionnaires based on TPB (McMillan & Conner 2003; Walker *et al.* 2004). Semantic differential scales help enhance the credibility of responses as this format facilitates asking several questions on the same component, e.g. the attitude towards respecting patients autonomy (DeVellis 2003, p.80). The respondents can be given different response categories in relation to different component in the same TPB questionnaire (McMillan & Conner 2003; Shankar *et al.* 2007; Walker *et al.* 2004) (Table 44).

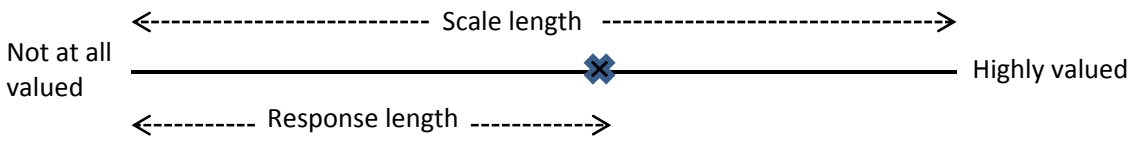
Table 44 - Response formats used commonly in TPB questionnaires

Format 1	Negative adjectival end (e.g. not valued at all)	1	2	3	4	5	6	7	Positive adjectival end (e.g. highly valued)
Format 2	Negative adjectival end (e.g. not valued at all)	-1	-2	-3	0	+1	+2	+3	Positive adjectival end (e.g. highly valued)

Semantic differential scales, however, may be more appropriate where there are few behaviours in a measure (DeVellis 2003, p.80). Using multiple adjectival scales for each of the 28 behavioural items used in this measure appeared impractical, as it would make the measure unacceptably lengthy. The length of a measure is a key determinant of both quality and quantity of responses especially in academic settings (Moss & Hendry 2002). On the other hand, in this study, a response to a TPB component for each item ultimately contributes to a particular latent variable (a facet). Therefore, a given TPB component in relation to a particular latent behaviour will actually be measured multiple times, which is the ultimate requirement to enhance the reliability and credibility of responses (Ajzen 2002; DeVellis 2003, pp.80-82; Francis *et al.* 2004). Therefore, a single question for each TPB component with a single bipolar adjectival scale, i.e. a visual analog scale, was used for this measure.

As visual analog scale is a variation of the semantic differential scale. However, unlike a semantic differential scale, in which the scale is ordinal, the scale of a visual analog is continuous; the respondent can place their response on a fixed-length (e.g. 10 cm) plain line, in which only the two ends are defined (e.g. left-end 'not at all valued' and right-end 'highly valued') (Figure 6).

Figure 6 – Example of a visual analog scale



Visual analog scales, however, have been used commonly for the measurements of feelings and emotions, e.g. pain, especially because of the practical difficulties associated with the manual measurement of response length (Couper *et al.* 2006). However, it can be used conveniently for surveying attitudes if a computer-based questionnaire is used; the calculation of the response length can be done automatically (Couper *et al.* 2006). A visual analog scale not only helps resolve the lengthening of the questionnaire but also helps the credibility of results as: it may be difficult for respondents to remember and reproduce exactly the same responses in repeated applications of the measure when visual analog scales are used; and a visual analog scale has the potential of identifying subtle differences in ratings within and between groups of respondents (Cicchetti *et al.* 2006; DeVellis 2003, pp.82-83). However, visual analog scales are associated with higher non-completion rates and missing data compared to other scale types (Couper *et al.* 2006). A visual analog scale was used as the response format for all three questions with possible protective strategies identified below to avoid non-completion and missing data. The right and left poles of the visual analog scale were determined in compliance with the TPB guidelines (Figure 7, p.1).

Measuring personal attitude: Personal attitude denotes the positive or negative evaluation of a behaviour by an individual (Ajzen 2002). As the value placed on a given

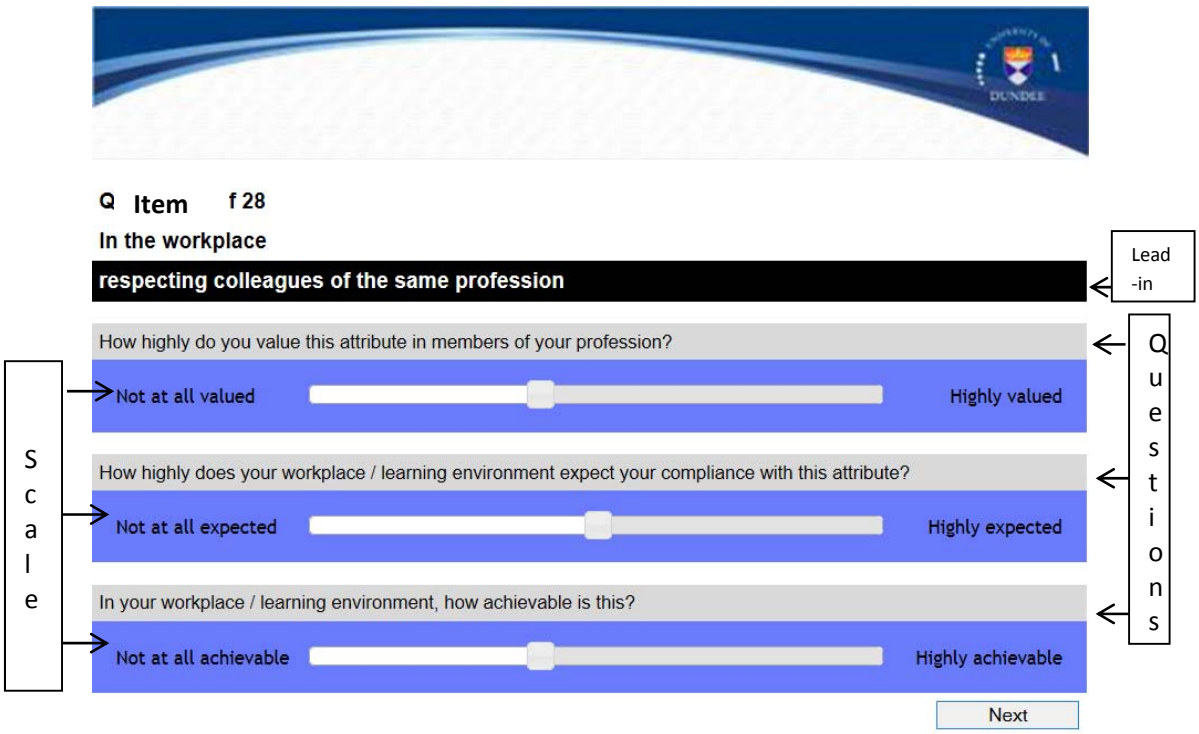
behaviour by an individual captures his or her overall evaluation (Francis *et al.* 2004), it was decided to use '*valuable'* as the 'adjective' in the rating scale.

Measuring subjective norm: 'Subjective norm' denotes the opinion of people, who usually interact with the respondent, about the behaviour concerned as perceived by the respondent (Ajzen 2002). In the context of measuring institutional culture, the 'expectations of the institution' (with regard to a given behaviour) was used to represent the 'people important to the respondent'. Institutional expectations are highly influential on how students perceive their professionalism behaviour (Monrouxe & Rees 2011). The degree of institutional *expectation*, therefore, was used in the rating scale.

Measuring perceived behavioural control: 'Perceived behavioural control' is a combination of self-efficacy (practical easiness and self-confidence of executing the behaviour) and self-control (degree of responsibility and influence of external factors) one has over a given behaviour (Ajzen 2002). Ideally, there should be questions on both self-efficacy and self-control in a strict measure of behavioural prediction (Ajzen 2002). However, as discussed above incorporating two questions would make the questionnaire lengthy. Therefore, it was decided to include a question that assessed the perceived *achievability* of a given behaviour, which would ultimately depict the final outcome of self-efficacy and self-control (Ajzen 2002).

The final version of an item (use of TACT in the lead, the three questions, and the visual analogy scale) is illustrated in Figure 7 using a screenshot of the actual measure as an example.

Figure 7 – A screenshot of the final measure depicting the structure of questioning in relation to any item



5.2. Design and delivery

The measure was developed into a web-based programme. The content and design of the programme was prepared by the author with the technical assistance of a professional computer programmer. The programme could be sent to the potential participants as an emailed URL link. (Please see p.202 for the full access details of the programme)

5.2.1. The format

The measure was designed to collect the responses of both students and members of the academic staff, and analyse and interpret results to illustrate the prevailing professionalism culture in a given medical school.

Several features were introduced into the programme to enhance its respondent-friendliness, credibility of responses and effectiveness and efficiency of obtaining, analysing and interpreting results.

- Although the items related to each of the facets clustered together in the measure, closely related items were not placed next to each other within the cluster. Each item was displayed sequentially one at a time on the screen and the respondents were not allowed to access any items, to which they had already responded. This helped minimise logical and proximity errors.
- Bogus items to detect inattentive responding were not included in this measure of professionalism culture. Unlike in traditional questionnaire surveys, there is no evidence to suggest that non-inclusion of bogus items causes unreliable results in a TPB based questionnaire (Francis *et al.* 2004). Although the mixing of positively worded lead-ins with negatively worded leads-ins and alternating the definition of the two ends of the scale (e.g. *highly valued* on the left end and *not at all valued* on the right for one item and the reverse for another) have been suggested, there is no supportive evidence to confirm the effectiveness of such strategies (Francis *et al* 2004). Using three questions in relation to each item may encourage attentive responding (Ajzen 2002). Furthermore, it has been demonstrated that visual analog scales can detect subtle differences even if the same item was given to the same group of respondents repeatedly as, unlike in Likert type scale, it provides no specific landmark on the rating scale (e.g. selecting response category 3 on a Likert scale on both occasions), which the

respondent can remember with certainty. Although inattentive responding cannot be totally eliminated, the use of three questions and a visual analog scale may help minimise this error to some extent.

- To respond, the participants could click at a particular point on the scale, or drag the indicator to the desired point. To overcome non-responding to items, which appears to be higher in visual analog scales than the conventional Likert-type scales, which can lead to missing values in the analysis (DeVellis 2003, pp.81-83), each of the three questions in a given item was highlighted once responded to. The respondent was not allowed to move to the next item until all three questions were highlighted. A 'pop-up' was used to advise the respondent on not-responded questions.

After completing all 28 items, the respondents were asked to provide certain demographic details; the university, age group, gender, student/staff status, if student, year of study, if staff main teaching involvement (undergraduate, postgraduate) and main professional involvement (clinical / non-clinical). An optional evaluation component was included to determine the reaction of respondents towards the measure. There were seven aspects to respond on a bipolar Likert-type scale (Table 45) with an area for open comments. As there were no such evaluations reported in the literature in relation to the measures of professionalism environment, the questions were developed by the author to evaluate the respondents' experience and the appropriateness of its content.

Table 45- Aspects of the measure included in the evaluation

Evaluation aspect	Purpose
The attributes included in the survey are representative of my understanding of professionalism	Gather evidence on face validity
The items were too hard to understand	Gather evidence of face validity and improve wording accordingly
The format and the presentation of the survey was attractive	Respondent-friendliness
The survey was too long	Respondent-friendliness
The results profile presented at the end portrayed my perception of the professionalism culture of this institution	Evidence for content validity
I enjoyed completing the survey	Respondent-friendliness

5.2.2. The scoring system

The visual analog scale in all three questions for each item carried a score between 0 – 10, as the conventional length of the scale (100 mm) is associated with the least error of measurement (Seymour *et al.* 1985). Each item was allocated equal weighting as there is no agreement on the contribution from each of the three components of TPB towards the prediction of intention (McMillan & Conner 2003; Shankar *et al.* 2007; Walker *et al.* 2004). The exact contribution from each component may be vital if the purpose of this study were to prove or disprove TPB. The TPB was used in this study only as the theoretical basis to improve the conceptuality and credibility of the measurement. Therefore, it was rational to allocate equal weight to each component. The scoring system used in this study is illustrated in Table 46.

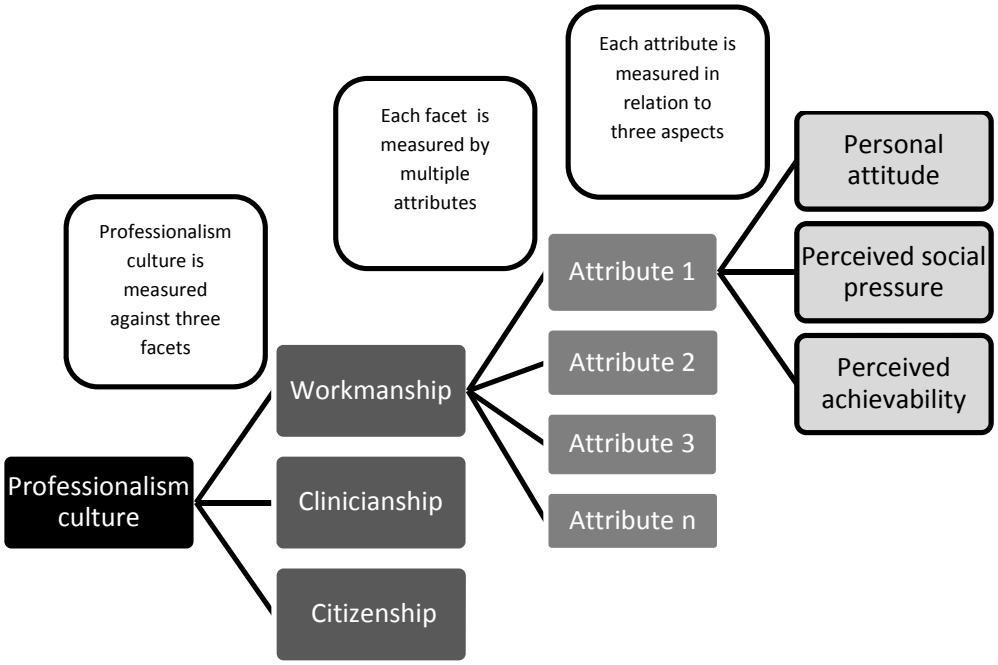
Table 46 – Scoring system used in the measure of institutional professionalism culture

	Number of items	Maximum possible total score for each question	No. questions	Maximum possible total score for each item	Final score (out of 10)
Working culture	11	10x11=110	3	110x3=330	(Respondent's score/330) x 10
Clinical culture	10	10x10=100	3	100x3=300	(Respondent's score/300) x 10
Social culture	7	10x7=70	3	70x3=210	(Respondent's score/70) x 10

5.2.3. Outputs, results and rationale for measurement components

The measure focused on the three facets of medical professionalism: the items under the facet of clinicianship represents the extent of patient-centeredness or practitioner-centeredness of behaviours during clinical encounters (clinical culture); the items under the facet of workmanship represents the extent of individual-orientation or team-orientation of behaviours in the workplace (working culture); and the items under the facet of citizenship represents the extent of self-centeredness or society-centeredness of personal behaviours (social culture). The responses gathered by the items under each facet determine the appropriate position within the respective extremes in a given institutional environment. For example, the overall average score of the ten items of clinicianship determines the position on the continuum between practitioner-centeredness and patient-centeredness. Therefore, each of the ten items acts as the test-item for patient-centered *versus* practitioner-centered behaviour. Both individual and collective feedback was provided on this basis. Figure 8 summarises the philosophical underpinning of the measure.

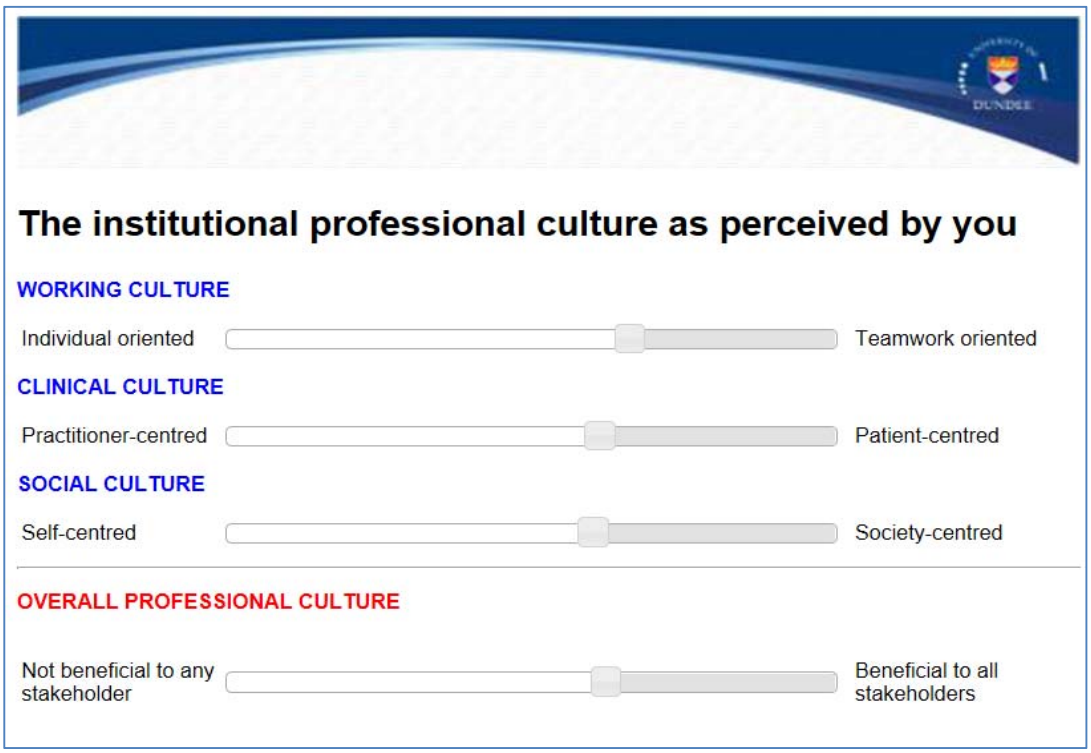
Figure 8 - Phylosophical underpinning of the measure



5.2.3.1. Individual feedback

Individual respondents were provided with an overall picture of the institutional professionalism culture based on his/her own responses after responding to all 28 items. (Figure 9)

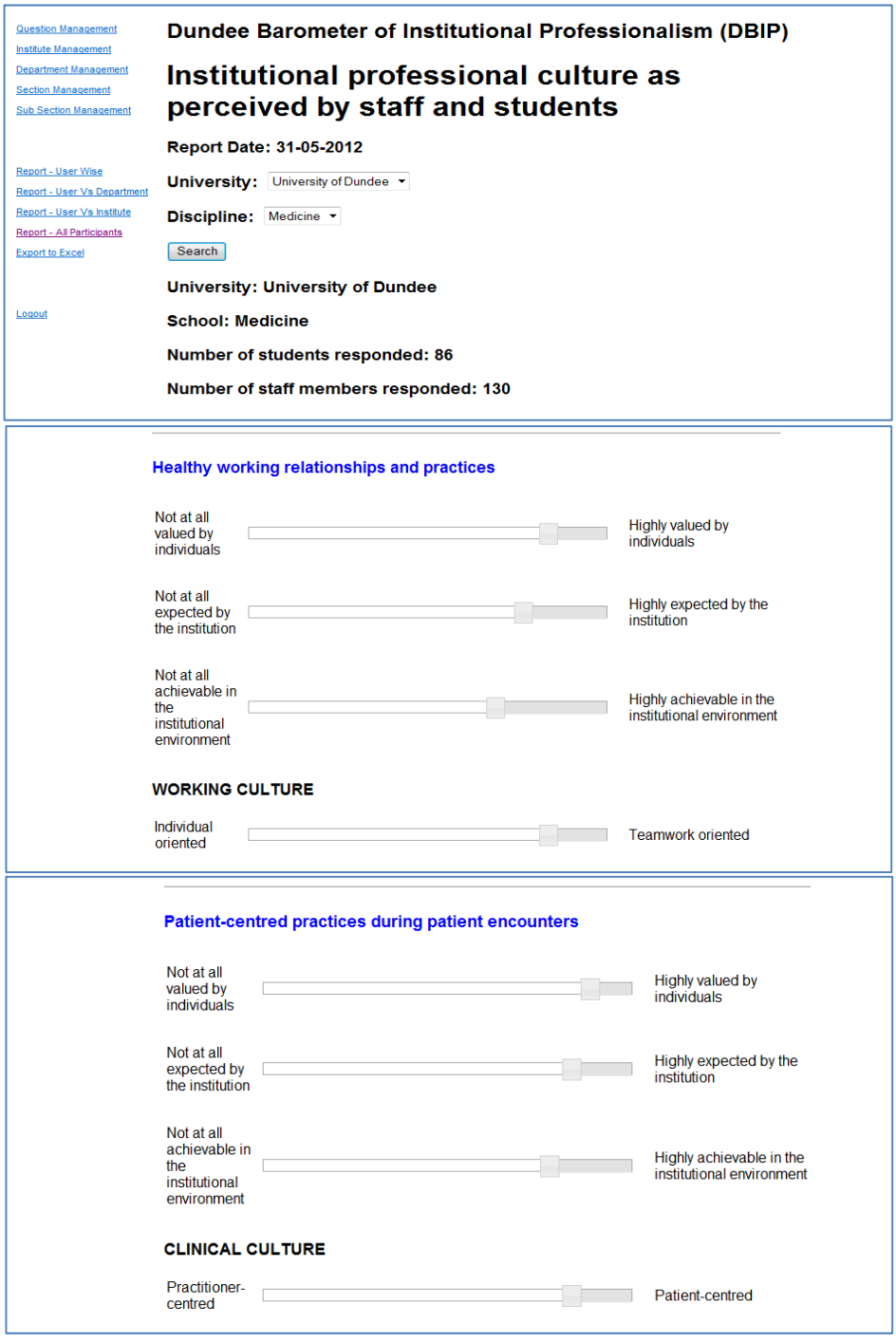
Figure 9 – Output generated for individual respondents based on one's own responses



5.2.3.2. Collective feedback

Immediately after each response, the programme automatically converts the responses to the measure to a graphical form which illustrates the professionalism culture of the institution as illustrated in Figure 10. This takes only a few clicks. The administrator of the measure in an institution was given password-protected access to overall results. In addition, the programme has the capacity of retrieving the individual responses to an Excel data sheet with ease that can be used for further statistical analysis using statistical software packages, e.g. Predictive Analytics SoftWare (PASW).

Figure 10 – Output generated for administrators based on overall responses




Socially responsible behaviours as a part of wider society
Not at all valued by individuals
Highly valued by individuals
Not at all expected by the institution
Highly expected by the institution
Not at all achievable in the institutional environment
Highly achievable in the institutional environment
SOCIAL CULTURE
Self-centred
Society-centred
OVERALL PROFESSIONAL CULTURE
Not beneficial to any stakeholder
Beneficial to all stakeholders

5.2.4. Final programme: the Dundee Barometer of Institutional Professionalism (DBIP)

The measure was named as the 'Dundee Barometer of Institutional Professionalism (DBIP)'. The development of the DBIP was an important part of this doctoral project. Hands-on experience of the DBIP can be obtained by accessing the measure using the following information.

To interact hands-on with the DBIP, please access:

http://chandratilake.com/DBIP/test

This link will directly take the user to a duplicate of the exact measure used in the first field test. No username or password is needed and, the user can interact with the programme freely.

(A printout of all screens of the online survey is included in Appendix VI)

To obtain an example of administrative output please access:

http://chandratilake.com/DBIP/test/admin

This link will take the user to an example of output template which can be obtained by the administrators through the DBIP cohort output mode. Please use the following username and password to access the output.

- User name: *dundeebiptest*

- Password: *20dbip11test*.

5.3. Discussion

This component of the study focused on the process of developing an online measure of the professionalism culture of medical schools by incorporating a theory-based response gathering system using the 28 items which represented the three facets of professionalism.

The unique feature of the DBIP compared to other measures of institutional professionalism culture discussed is its use of a theory-based approached to measurement. The assessment of professionalism has been frequently criticised for the lack of theoretical approach, and the need for a theoretical underpinning for such measures has been repeatedly emphasised (Archer *et al.* 2008; Jha *et al.* 2007; Rees & Knight 2007). The DBIP, however, does not prove or disprove the theory, on which it was based (TPB). Rather it uses the most suitable theoretical underpinning to enhance the credibility of responses; a strategy which is recommended for developing all surveys (DeVellis 2003, p.8). A UK study reported that the TPB questionnaire format minimised the possibility of selecting a socially acceptable response regardless of one's own opinion / experience (Armitage & Conner 1999), which is identified as a major source of bias and a drawback in professionalism surveys (Baldwin & Self 2006, p.102). TPB, which underpinned the response gathering system of the DBIP, has been recommended as appropriate for measuring and predicting professional behaviours (Archer *et al.* 2008; Rees & Knight 2007). However, although the perceptions gathered through this system tend to be considerably more credible than those from a simple rating scale, some level of

disparity between perceived reality and the reality may still exist; TPB does not predict actual behaviour with 100% certainty (Armitage & Conner 2001). The use of TPB encompasses the concept of 'institutional culture' fully by not only focusing on attitudes of people or behaviours but also considering social norms and pressures (Hudelson 2004).

Visual analog scales were not used by any of the measures of professionalism culture discussed in the literature review, and are used infrequently in surveys (Clark & Watson 1995; Couper *et al.* 2006). Amidst the evidence to support visual analog scales being more discriminatory than conventional rating scales, the scarcity of their use in surveys may be attributable to practical difficulties (Couper *et al.* 2006). The use of a visual analog scale may have helped overcome key deficiencies of Likert-type scales in measuring perceptions, e.g. inattentive responding, logical error, non-discriminating results with repeated measures, as anticipated. However, there is no comparative means of confirming such an assumption in this study.

The use of an online mode for the DBIP has both advantages and disadvantages. The approach provided infrastructure for using a visual analog as a rating scale, which would have been almost impractical in a multi-item paper-based survey. It also may have helped overcome logical and proximity errors as one item was presented at a given time. This strategy, together with the visual analog scale, may have had a synergistic effect against the potential deficiencies of self-rating mentioned above. Missing responses are a common problem associated with surveys and may cause considerable bias (de Leeuw 2001). Although the quality of questions and questionnaire design play a part, there is a

higher probability of missing responses in paper-based surveys (de Leeuw 2001). The web-based approach, which is known to yield comparatively fewer missing data in surveys (de Leeuw 2001), together with making the responses compulsory for every question via the software programme, may have ensured the final results contained no missing responses. The majority of the measures of professionalism culture discussed in the literature review were pen-and-paper based except for the polyprofessionalism inventory (Roff *et al.* 2011) and moral distress survey (Wiggleton *et al.* 2010). The Professionalism climate instrument has been delivered both in paper and email versions (Quaintance *et al.* 2008). Although no author has discussed the reasons for using or not using pencil-and-paper approach, the primary obstacle may have been the lack of wide-accessibility to computers during the respective time periods. However, Golsing *et al* (2004) claim that, in addition to the limited accessibility to computers, the misconceptions such as 'internet users differ from nonusers', 'internet users are unmotivated', 'internet-based survey findings differ from those obtained with other methods', all of which were proven to be untrue in meta-analytic and comparative studies, may have caused the inclination towards pen-and-paper mode. The analysis and interpretation of results was made easy with nature of the software used in this study, which generated individual and collective feedback, reducing the workload of the administrators. This is particularly advantageous, as there has been a constant worry among medical educationalists about the resource-intensiveness of professionalism assessments and the scarcity of providing feedback to the assessee / respondent (van Mook *et al.* 2009b). However, the web-based delivery mode may have

had a negative effect on the response rate which is discussed below under the findings of the first field-test.

Chapter 6 – The first field-test of the measure of professionalism culture

This chapter discusses the process and outcome of first field-test of the DBIP.

6.1. Methodology

6.1.1. Ethics

Ethics approval for the first field test was obtained from the University Research Ethics Committee (UREC) of the University of Dundee UK (Reference No: UREC 11023). Permission and support was obtained from the Medical School Board, Dean and Teaching Dean of the Medical School, University of Dundee to conduct the study.

6.1.2. Study design

The first field test was conducted as a cross sectional survey.

6.1.3. Setting and context

The study was conducted between 15 September and 31 November 2011 in Dundee Medical School. Many developers of measures on professionalism culture have used the institution, to which they are attached, for the first field test to determine the psychometric properties (Hojat *et al.* 2001; Hojat *et al.* 2009a; Kalet & Steven 2004; Quaintance *et al.* 2008; Roff *et al.* 2011; Wiggleton *et al.* 2010).

6.1.4. Sample

Fourth and final year medical students of the Dundee Medical School (n=321) were invited to participate in the study. The two cohorts were chosen on the advice of the Teaching Dean, who reviewed the items and recommended that, given the nature of clinical

exposure in the Dundee curriculum, only the students in these years would be able to credibly respond to the measure. The faculty was represented by the academic staff of the Dundee Medical School and the extended faculty (NHS Clinicians who hold honorary contracts with the University of Dundee). The medical school office sent the email invitation to participate in the first field test to a pool of 700 members of the University and extended faculty members.

6.1.5. Data collection

The DBIP as a URL web-link was sent to potential participants through the medical school office with a brief description. The email also contained a message from the Teaching Dean, which encouraged participation in the DBIP survey. Such endorsements from an authority appear to increase the response rate for web-surveys among university students (Joinson *et al.* 2007; Porter & Whitecomb 2003). The emails, however, were not personalised. It is recommended to send personalised email invitations to university students in order to improve response rates (Mitra *et al.* 2008). However the literature evidence that it actually enhances the response rates in practice is inconclusive (Porter & Whitecomb 2003) especially when the information collected is sensitive and personalisation may be seen to potentially compromise anonymity (Joinson *et al.* 2007). Two reminders were sent in the first and second weeks of the launch, the points at which the reminders would be most effective (Deutskens *et al.* 2004). More reminders were avoided as ethics permission was granted only for two reminders and there was no evidence to suggest that more than two reminders make any significant difference to the response rate. Time periods, in which assessments were scheduled and clinical rotations

started, were avoided as advised by the Dean of Teaching, to minimise the disturbance to potential respondents.

6.1.6. Data analysis

Individual responses were transferred to a Predictive Analytic Software (PASW®, formally known as SPSS) file through an Excel data sheet generated by the DBIP programme itself. All statistical analyses were carried out using PASW® version 20.

The demographic characteristics of respondents were analysed using descriptive statistics. The Cronbach alpha (Cronbach 1951), a measure of internal consistency based on the classical test theory, was used in determining the internal structure of the DBIP at facet and the measure level. This has frequently been used for the purpose by developers of measures (Arnold *et al.* 1998; Hojat *et al.* 2001; Hojat *et al.* 2003b; Quaintance *et al.* 2008; Thrush *et al.* 2011).

Average scores of questions, items, facets and overall measure levels were computed for statistical analysis. Table 47 explains the method used in computing these average scores.

Table 47 – Methods of calculating average scores for statistical analysis

Average score	Method of calculation
Item scores	The average for (1) personal attitude, (2) institutional expectation and (3) achievability scores for each item.
Facet attitude score	The average for personal attitude scores of all items of a given facet.
Facet expectation score	The average for expectation scores of all items of a given facet
Facet achievability score	The average for achievability scores of all items of a given facet
Facet total score	The average of (b), (c) and (d) above for each facet
Total average score	The average of facet total scores for the three facets

The frequency distribution of the facet and total scores for normalcy was assessed using histograms and Kolmogorov-Smirnov (K-S) test; histograms demonstrate the deviation and K-S tests demonstrate whether the deviations are statistical significant (Field 2009, pp.145-147). As the analysis indicated that the scores appeared to be heavily towards the right, i.e. negatively skewed, and deviations were statistically significant, inferences were made using non-parametric statistical methods.

The mean scores of different demographic groups for the facets and the entire measure were compared only when the groups concerned had adequate number of respondents, as statistical inferences made with insufficient subjects in comparing groups may be inaccurate and misleading (Field 2009, p.58). The demographic analysis of respondents revealed that the only comparison possible was between students and faculty, and the gender differences among the faculty. These comparisons were carried out using the Mann Whitney U test, which is the non-parametric equivalent of the independent sample *t*-test (Field 2009, pp.540-551). The correlations between the three facets were examined using Spearman correlation (ρ) to determine their relationship to each other (Field 2009 ,pp.177-179).

The responses to evaluation questions with Likert-type rating scales were analysed using descriptive and non-parametric inferential statistics as the data was ordinal (Field 2009, pp.540-551). Accordingly, the responses of students and faculty were compared using the Mann Whitney U test. The open comments made by the respondents were thematically analysed.

In instances where the statistical significance was important, it was determined with a 95% confidence interval (p <0.05) (unless stated otherwise), which is deemed to be acceptable for a measure such as the DBIP (Field 2009, p.51).

Complying with measures of professionalism culture (Arnold *et al.* 1998; Hojat *et al.* 2001; Hojat *et al.* 2003b; Quaintance *et al.* 2008; Thrush *et al.* 2011), which have been validated and used widely, it was decided to reduce the items if possible, to a more practicable number without compromising the psychometric rigour. As expressed in the evaluation (Table 55, p.223), the length of the surveys was a 'concern' for the respondents. Principal Component Analysis (PCA) is a frequently used statistical method for item reduction (Clark & Watson 1995). Therefore, a principal component analysis, specifying three components (as the DBIP contained three pre-determined facets) and an orthogonal rotation was conducted to identify a suitable number of items with the highest representation of and relationship to each facet (DeVellis 2003, pp.102-137).

The items selected from the PCA were subjected to generalisability analysis. This analysis is based on the Generalisability (G) Theory (Cronbach *et al.* 1963). The resultant, G coefficient is the ratio of universal score variance to observed score variance, which is analogous to alpha coefficient (DeVellis 2003 ,pp.44-47). However, a G study also produces an analysis of the error component; if the majority of error is explained by respondents, it can be anticipated that the measure is able to detect the variations in responses with high accuracy (DeVellis 2003 ,pp.44-47). A Decision (D) study based on G theory predicts the reliability of measures with different combinations of item / questions

/ domain (Gaessaroli 2006). A D study was conducted on the selected items to examine the psychometric rigiour of the final measure with reduced number of items.

To examine the meaningfulness of high reliabilities observed, a correlation analysis of the selected items to each other was conducted. It was anticipated that the correlation of items within a particular facet should be greater than that of between facets (Clark & Watson 1995).

6.2. Findings

The findings of the first field-test are presented under several headings.

- Demographic characteristics of respondents
- Internal structure and statistical characteristics of the measure
- Item, domain and total scores of the measure
- Comparison of demographic groups
- Feedback of respondents towards the measure
- Factor analysis and item reduction
- Psychometric analysis of final version of the measure

6.2.1. Demographic characteristics of respondents

The DBIP was responded to by 212 members of the Dundee Medical School; 127 members of the staff (response rate 19%), and 85 students (response rate 25%). Their demographic characteristics are shown in Table 48.

Table 48 - Demographic characteristics of the respondents to the DBIP first field test

Stakeholder group	Demographic characteristics		Number	%	Cumulative %
Staff (n=127)	Gender	Male	75	59.1	59.1
		Female	49	38.6	97.6
		Prefer not to say	3	2.4	100.0
	Age group	19 - 25	2	1.6	1.6
		26 - 35	24	18.9	20.5
		36 - 60	97	76.4	96.9
		>60	4	3.1	100.0
	Professional engagement	Mainly clinical	104	81.9	81.9
		Mainly non-clinical	23	18.1	100.0
Students (n=85)	Gender	Male	23	27.1	27.1
		Female	61	71.8	98.8
		Prefer not to say	1	1.2	100.0
	Age group	19 - 25	75	88.2	88.2
		26 - 35	8	9.4	97.6
		36 - 60	2	2.4	100.0
	Year of study	Fourth year	42	49.4	49.4
		Fifth year	43	50.6	100.0

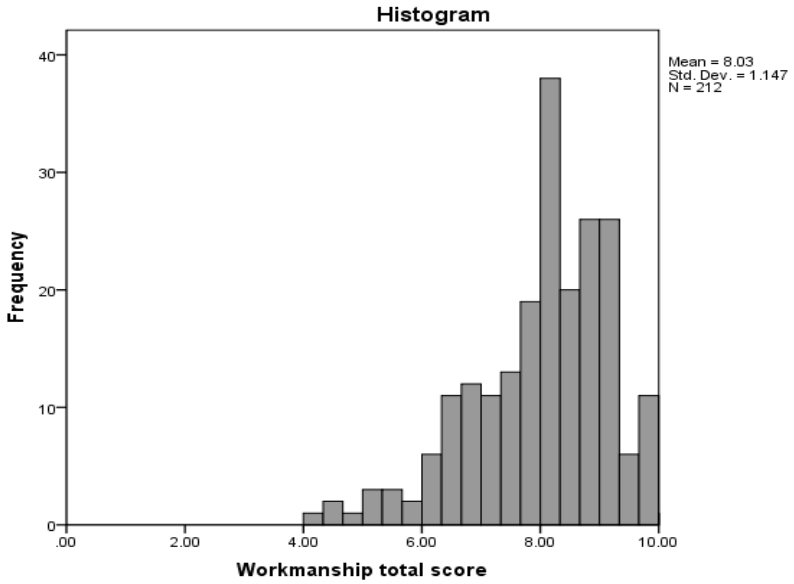
The representativeness of the targeted populations by the respondents in terms of demographic characteristics such as age group, gender (for both students and staff) and professional engagement (for staff) could not be determined, as such information was not readily available with the Dundee Medical School Office.

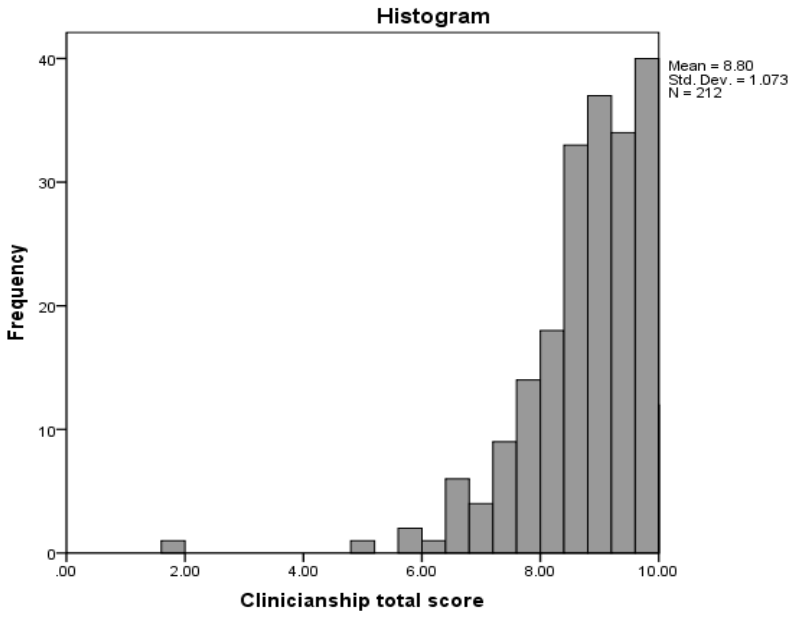
6.2.2. Internal structure and statistical characteristics of the measure

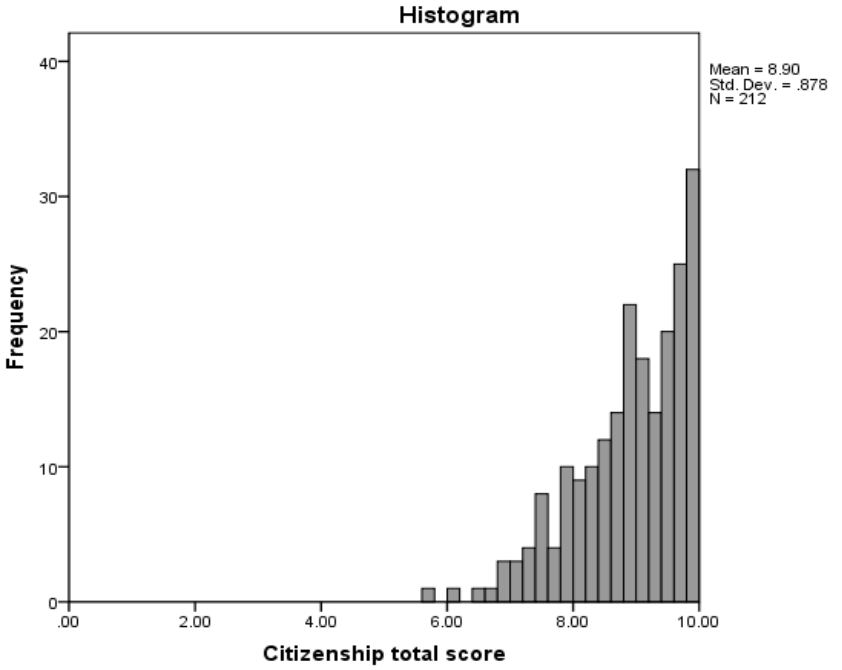
The internal consistency of 28 items, as measured by the Cronbach alpha was 0.953. This indicated that all 28 items were strongly interrelated (DeVellis 2003, pp.27-29).

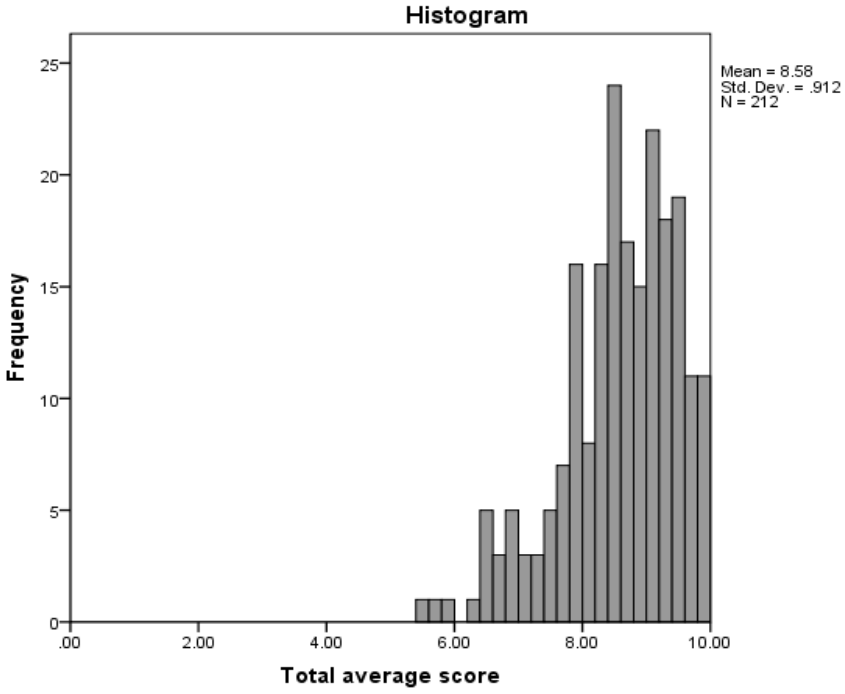
The frequency histograms clearly illustrated that the scores were grossly skewed to the right (Figure 11).

Figure 11 – Frequency distributions of facet and total scores in the first field test






Histogram
Mean = 8.90
Std. Dev. = .878
N = 212
Frequency
40
30
20
10
0
.00
2.00
4.00
6.00
8.00
10.00
Citizenship total score


Histogram
Mean = 8.58
Std. Dev. = .912
N = 212
Frequency
25
20
15
10
5
0
.00
2.00
4.00
6.00
8.00
10.00
Total average score

K-S test indicated that the deviation of all scores from normal distribution is significant, i.e. the scores were not normally distributed (Table 49) and it is more appropriate to use non-parametric statistics (Field 2009, pp.145-148).

Table 49 - Significance of normal distribution as measured by Kolmogorov-Smirnova (K-S) test

(degree of freedom =212)	Statistic	*p*
Workmanship total score	0.101	0.000
Clinicianship total score	0.131	0.000
Citizenship total score	0.106	0.000
Total average score	0.078	0.003

6.2.3. Item, domain and total scores of the measure

The scores of all items (therefore, all facets and institutional culture in general) inclined heavily towards the *professional* rather than *unprofessional* ends of the respective scoring scales, i.e. presence of positive personal attitudes, high institutional expectations and a conducive environment with regards to all items. This suggested the existence of an institutional professionalism culture which is beneficial to all. (Figure 12, Table 50 and Table 51)

Figure 12 – The professional culture of Dundee Medical School as illustrated by the Dundee Barometer of Institutional Professionalism

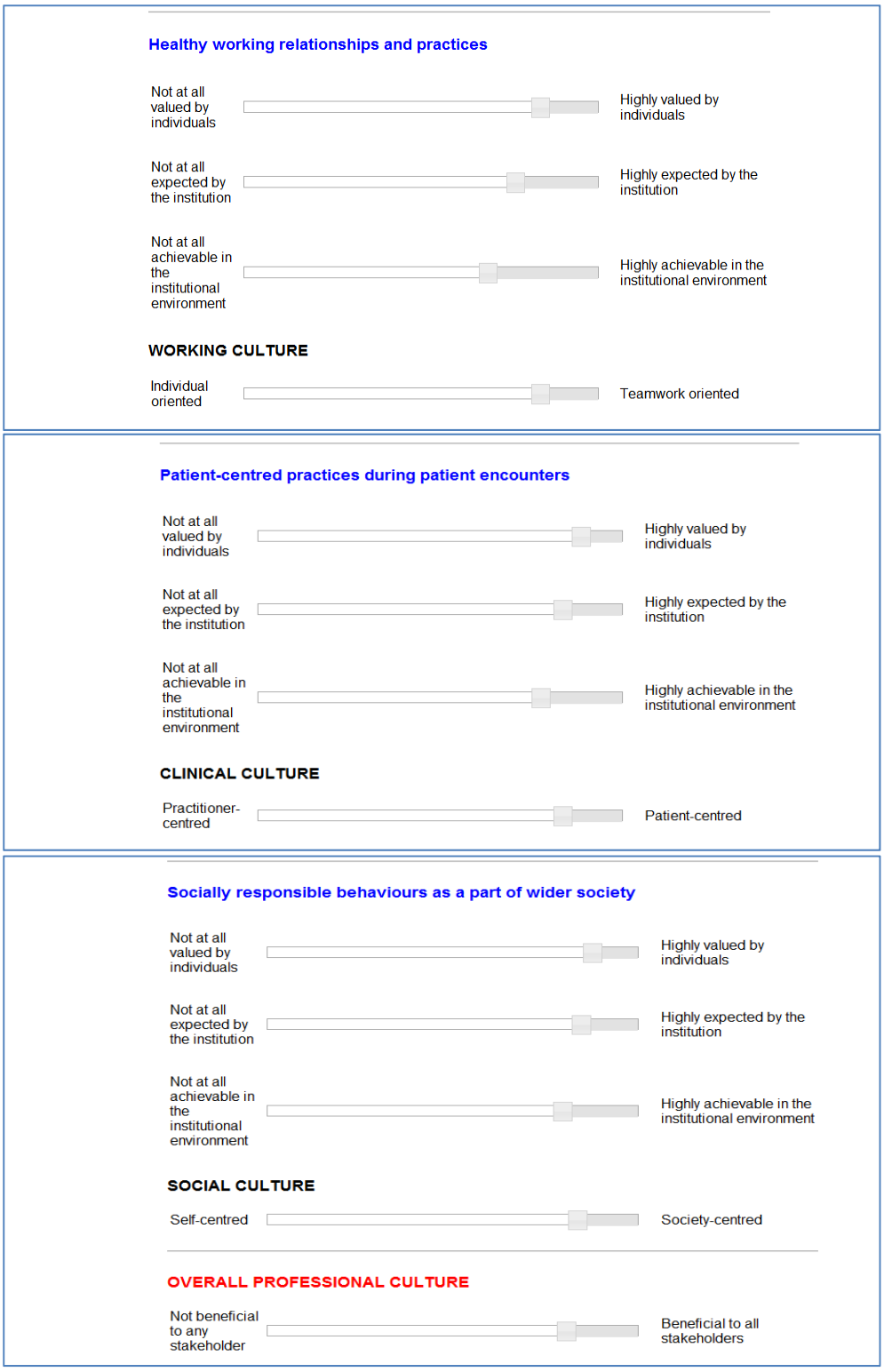


Table 50 – Average scores for each of the 28 items of the DBIP

Item (Sorted according to descending order of average total scores)	Average scores			
	Attitude	Expectation	Achievability	Total
Functioning according to the law	9.46	9.48	9.17	9.37
Behaving honestly and with integrity	9.64	9.11	8.86	9.20
Acting in a responsible fashion towards patients	9.50	9.16	8.89	9.18
Treating patients fairly and without prejudice	9.51	9.11	8.73	9.12
Respecting patients' confidentiality and privacy	9.58	9.33	8.38	9.10
Avoiding substance or alcohol misuse	9.06	8.93	8.74	8.91
Behaving in a reliable and dependable way	9.38	8.91	8.42	8.90
Being sound in judgment and in decision making	9.44	8.83	8.27	8.85
Communicating with patients in a clear and effective manner	9.42	8.83	8.10	8.78
Adhering to professional rules and regulations	9.00	9.04	8.26	8.77
Being accountable for one's actions	9.41	8.67	8.20	8.76
Showing compassion towards patients	9.41	8.51	8.30	8.74
Being empathetic when caring for patients	9.33	8.48	8.34	8.72
Treating colleagues of the same profession fairly and without prejudice	9.32	8.60	8.08	8.67
Respecting patients' autonomy	9.17	8.65	8.03	8.62
Treating other healthcare professionals fairly and without prejudice	9.34	8.46	8.01	8.60
Working well as a member of a team	9.26	8.56	7.63	8.48
Taking a dedicated approach to work	9.03	8.47	7.88	8.46
Being attentive to the needs of patients	9.23	8.57	7.48	8.42
Providing advice to patients when required	9.05	8.54	7.65	8.41
Respecting colleagues of the same profession	9.00	8.11	7.89	8.33
Being aware of own limitations	9.05	7.97	7.33	8.11
Not using professional status for personal gain	8.79	7.61	7.56	7.98
Reflecting on your actions with a view to self-improvement	8.54	8.05	6.79	7.79
Making effective use of the resources available	8.35	8.16	6.44	7.65
Having a positive attitude towards professional development	8.45	7.92	6.56	7.64
Being receptive to constructive criticism	8.58	7.78	6.54	7.63
Being able to manage situations where there is a conflict of interest	8.43	7.49	6.24	7.39

Although the difference of total score between facets was very small the highest and lowest scores were recorded against the citizenship and workmanship facets respectively.

Across all three domains, the scores of personal attitude remained the highest and the scores of achievability remained the lowest. (Table 51)

Table 51 - Average scores for the domains of attitude, expectation and achievability in each facet

Facet	Total sub-domain scores			Total domain score
	Attitude	Expectation	Achievability	
Citizenship	9.29	8.93	8.48	8.9
Clinicianship	9.36	8.81	8.23	8.8
Workmanship	8.83	8.06	7.19	8.03

The correlation between average scores for the three TPB components across the three facets were moderate (p <0.05), which suggests that the three questions measure distinct but related aspects of each behaviour (Table 52).

Table 52 - Correlation between average scores for the domains attitude, expectation and achievability for each of facets

ρ = Spearman correlation *sig.* = Significance (<0.01)		Workmanship: attitude	Workmanship: expectation	Workmanship: achievability
Workmanship: attitude	ρ	1.000	0.670	0.486
	Sig.	-	0.000	0.000
Workmanship: expectation	ρ	0.670	1.000	0.566
	Sig.	0.000	-	0.000
Workmanship: achievability	ρ	0.486	0.566	1.000
	Sig.	0.000	0.000	-
		Clinicianship: attitude	Clinicianship: expectation	Clinicianship: achievability
Clinicianship: attitude score	ρ	1.000	0.725	0.531
	Sig.	-	0.000	0.000
Clinicianship: expectation score	ρ	0.725	1.000	0.625
	Sig.	0.000		0.000
Clinicianship: achievability score	ρ	0.531	0.625	1.000
	Sig.	0.000	0.000	-
		Citizenship: attitude	Citizenship: expectation	Citizenship: achievability
Citizenship: attitude score	ρ	1.000	0.710	0.534
	Sig.	-	0.000	0.000
Citizenship: expectation score	ρ	0.710	1.000	0.579
	Sig.	0.000	-	0.000
Citizenship: achievability score	ρ	0.534	0.579	1.000
	Sig.	0.000	0.000	-

6.2.4. Comparison of demographic groups

The comparisons were conducted using the Mann Whitney U test. This test ranks the scores and the differences between ranks are used to determine the statistical significance; the higher the rank the higher the score and *vice versa* (Field 2009, pp.540-551).

Several differences were observed between student and faculty scores. The item scores and facets scores of students were higher than those of the faculty, and in most instances, the differences were statistically significant (Table 53). However, for the avoidance of substance and alcohol misuse students scored significantly lower than their teachers.

Unlike students and faculty, no general pattern between male and female faculty was observed for item or facet scores. However, as indicated by certain total items scores and fact-level attitudes scores, the female faculty appears to have a significantly positive attitude towards the relationship between doctors and co-workers and doctors and patients compared to their male counterparts (Table 54).

The comparison of other demographic sub-groups within faculty and students were abandoned as the numbers representing each subgroup were inadequate for a meaningful statistical comparison.

Table 53 - Comparison of students and faculty responses to each item

(Faculty n=127 and students n=85. Significant statistical differences (>0.05) are indicated in bold)

Item / facet	Median		Mean rank		
	staff	student	staff	student	*p*
Respecting colleagues of the same profession	8.00	8.67	93.78	125.51	**0.000**
Reflecting on your actions with a view to self-improvement	7.33	8.67	91.34	129.15	**0.000**
Having a positive attitude towards professional development	7.33	8.67	88.61	133.24	**0.000**
Working well as a member of a team	8.33	9.33	90.92	129.78	**0.000**
Being aware of own limitations[1]	8.33	8.33	99.50	116.96	**0.041**
Being receptive to constructive criticism	7.33	8.33	96.23	121.84	**0.003**
Treating other healthcare professionals fairly and without prejudice	9.00	9.33	98.12	119.02	**0.014**
Being able to manage situations where there is a conflict of interest	7.33	7.67	100.82	114.99	0.099
Treating colleagues of the same profession fairly and without prejudice	9.00	9.33	95.22	123.35	**0.001**
Towards not using professional status for personal gain	8.33	8.33	102.86	111.94	0.288
Making effective use of the resources available	7.67	8.00	100.92	114.84	0.104
Respecting patients' autonomy	8.67	9.33	95.54	122.88	**0.001**
Behaving in a reliable and dependable way	9.00	9.67	98.35	118.68	**0.016**
Communicating with patients in a clear and effective manner	9.00	9.33	99.91	116.34	0.054
Acting in a responsible fashion towards patients	9.33	9.67	98.60	118.30	**0.016**
Being attentive to the needs of patients	8.67	9.00	105.02	108.71	0.666
Providing advice to patients when required	8.67	9.00	103.91	110.36	0.450
Showing compassion towards patients	9.00	9.33	101.49	113.98	0.140
Respecting patients' confidentiality and privacy	9.33	9.67	99.85	116.44	**0.047**
Treating patients fairly and without prejudice	9.33	9.67	103.63	110.79	0.390
Being empathetic when caring for patients	9.00	9.33	97.68	119.68	**0.010**
Being accountable for one's actions	8.67	9.33	94.08	125.06	**0.000**
Behaving honestly and with integrity	9.33	9.67	100.92	114.84	0.094
Adhering to professional rules and regulations	9.00	9.33	103.69	110.71	0.406
Functioning according to the law	10.00	10.00	106.11	107.08	0.901
Avoiding substance or alcohol misuse	9.67	9.33	116.60	91.41	**0.002**
Being sound in judgment and in decision making	9.00	9.33	104.93	108.85	0.645
Taking a dedicated approach to work	8.67	9.00	103.02	111.71	0.309
Workmanship attitude score	9.00	9.09	101.81	113.50	0.174
Workmanship expectation score	8.09	8.55	92.80	126.98	**0.000**
Workmanship achievability score	7.18	7.64	95.08	123.56	**0.001**
Workmanship total score	8.00	8.42	93.49	125.94	**0.000**
Clinicianship attitude score	9.60	9.70	104.89	108.91	0.632
Clinicianship expectation score	9.00	9.40	100.41	115.61	0.076
Clinicianship achievability score	8.30	8.70	99.77	116.56	0.051
Clinicianship total score	8.80	9.17	99.58	116.84	**0.045**
Citizenship attitude score	9.57	9.57	110.36	100.74	0.254
Citizenship expectation score	9.14	9.57	101.61	113.80	0.154
Citizenship achievability score	8.57	8.86	105.34	108.23	0.737
Citizenship total score	9.00	9.10	104.77	109.08	0.616
Total average score	8.60	8.93	97.76	119.56	**0.011**

[1] Occasionally, the difference between two groups can be statistically significant with median being the same for both groups.

Table 54 – Gender differences in item and facet scores observed among the faculty

(Males n=75 and females n=49. Significant statistical differences (>0.05) are indicated in bold)

	Median		Mean rank		
	Male	Female	Male	Female	*p*
Respecting colleagues of the same profession	8.00	8.00	63.15	61.51	0.803
Reflecting on your actions with a view to self-improvement	7.33	7.67	58.95	67.94	0.172
Having a positive attitude towards professional development	7.33	7.67	60.49	65.58	0.439
Working well as a member of a team	8.33	8.67	61.31	64.32	0.648
Being aware of own limitations	8.33	8.33	61.47	64.08	0.691
Being receptive to constructive criticism	7.67	7.33	61.40	64.18	0.673
Treating other healthcare professionals fairly and without prejudice	9.00	9.00	62.89	61.90	0.879
Being able to manage situations where there is a conflict of interest	7.67	7.33	65.61	57.73	0.232
Treating colleagues of the same profession fairly and without prejudice	9.00	8.67	63.42	61.09	0.723
Towards not using professional status for personal gain	8.33	8.67	60.83	65.06	0.519
Making effective use of the resources available	7.67	7.67	61.36	64.24	0.661
Respecting patients' autonomy	8.33	9.33	56.77	71.28	**0.027**
Behaving in a reliable and dependable way	9.00	9.33	56.24	72.08	**0.015**
Communicating with patients in a clear and effective manner	8.67	9.33	56.43	71.80	**0.019**
Acting in a responsible fashion towards patients	9.33	9.67	59.45	67.16	0.226
Being attentive to the needs of patients	8.33	9.00	59.17	67.60	0.199
Providing advice to patients when required	8.67	9.00	59.58	66.97	0.260
Showing compassion towards patients	9.00	9.00	61.80	63.57	0.785
Respecting patients' confidentiality and privacy	9.33	9.67	58.07	69.29	0.081
Treating patients fairly and without prejudice	9.00	9.67	58.39	68.79	0.106
Being empathetic when caring for patients	8.67	9.00	59.14	67.64	0.194
Being accountable for one's actions	8.67	8.67	60.73	65.20	0.495
Behaving honestly and with integrity	9.33	9.67	59.63	66.89	0.258
Adhering to professional rules and regulations	9.00	9.33	60.59	65.43	0.455
Functioning according to the law	9.67	10.00	59.22	67.52	0.171
Avoiding substance or alcohol misuse	9.33	9.67	61.30	64.34	0.631
Being sound in judgment and in decision making	9.00	9.00	62.55	62.43	0.986
Taking a dedicated approach to work	8.67	9.00	60.56	65.47	0.454
Workmanship attitude score	8.73	9.18	56.89	71.08	**0.031**
Workmanship expectation score	8.09	8.09	62.02	63.23	0.854
Workmanship achievability score	7.36	6.91	63.89	60.37	0.593
Workmanship total score	8.06	8.06	62.38	62.68	0.963
Clinicianship attitude score	9.40	9.90	55.21	73.66	**0.004**
Clinicianship expectation score	8.90	9.30	59.06	67.77	0.185
Clinicianship achievability score	8.10	8.80	58.91	68.00	0.168
Clinicianship total score	8.67	9.23	57.93	69.50	0.079
Citizenship attitude score	9.43	9.86	57.78	69.72	0.063
Citizenship expectation score	9.14	9.29	59.91	66.47	0.317
Citizenship achievability score	8.57	8.57	64.00	60.20	0.565
Citizenship total score	8.95	9.10	60.71	65.23	0.493
Total average score	8.59	8.68	60.33	65.82	0.406

6.2.5. Feedback of respondents towards the measure

One hundred and eighty three (112 staff members and 72 students) of 212 respondents completed the evaluation component of the DBIP. The feedback was generally positive. However, the feedback on the length of the survey and enjoyment of completing it were inconclusive. The responses of students and staff to the evaluation questions were not significantly different. (Table 55)

Table 55 - Feedback on the survey design and content received from the field-test participants (staff and students of the Medical School)

Questions	Number of responses in response categories (%)					Median rating			
	Strongly disagree	Disagree	Unsure	Agree	Strongly agree	Overall	staff	students	*p*
The attributes included in the survey are representative of my understanding of professionalism	2 (0.9)	6 (2.8)	20 (9.4)	114 (53.8)	38 (17.9)	4.00	4.00	4.00	0.45
The items were too hard to understand	45 (21.2)	91 (42.9)	21 (9.9)	19 (9.0)	4 (1.9)	2.00	2.00	2.00	0.76
The format and the presentation of the survey was attractive	7 (3.3)	23 (10.8)	31 (14.6)	102 (48.1)	17 (8.0)	4.00	4.00	4.00	0.68
The survey was too long	7 (3.3)	66 (31.1)	31 (14.6)	55 (25.9)	21 (9.9)	3.00	3.00	3.00	0.22
The results profile presented at the end portrayed my perception of the professionalism culture of this institution	9 (4.2)	9 (4.2)	49 (23.1)	92 (43.4)	21 (9.9)	4.00	4.00	4.00	0.08
I enjoyed completing the survey	10 (4.7)	39 (18.4)	59 (27.8)	65 (30.7)	6 (2.8)	3.00	3.00	3.00	0.91

The free comments in the DBIP evaluation focused mainly on the third question of every item, technical issues, content and considering non-academic staff as potential respondents (Table 56).

Table 56 – Qualitative comments of respondents on the Dundee Barometer of Institutional Professionalism

Theme	Comments
About the third questions	*'The questions especially the third part of each question is open to debate. What do you mean by achievable - achievable by me, the organisation, the team? Given this the answers to the questions are not representative of what I think.'*
	'The wording of each third item is difficult - things might well be achievable, but on an everyday basis, they don't happen. So my answers to the third items reflect more what actually happens than what is achievable in theory.'
	'The 'wider society' section was a little confusing - incorporating ideas about wider society into institutional culture.'
	'May have answered the last question of each section with a different understanding than intended. Sometimes I think there is a difference in the reality of the actual professionalism delivered because of other factors e.g. communicating effectively with patients I think is important.'
	'In the last questions re wider society I did not feel answering re the workplace made sense, hence put all responses in middle as couldn't opt out.'
Technical aspects	*'Very slow going from one question to the next, may result in poor response!'*
	'Sliding scale may have benefited from markers at 1/4, 1/2 etc.'
Content	*'Perhaps avoid jargon related to professionalism and put the attributes more into context'*
	'It seemed a bit repetitive. Also I wasn't sure that the survey was that discriminatory; nearly all of my responses were towards the 8, 9 or 10/10 agreement.'
	'Perhaps a free text box at the end for explanations of why there might be a discrepancy between individual and institutional ideals of professionalism and why they might be hard to achieve.'
	'I don't think that you have really identified 'institutions' well enough (maybe I read it too quickly). I have answered in terms of my clinical practice not my academic institution. So anything relating to my workplace is for a little general practice with big organisational problems.'
Accessing non-clinical staff	*'As a wholly non-clinical staff member, the questions on dealing with patients were not relevant. Because there was no option to indicate this, I had to answer neutrally for them all - which has doubtless skewed my overall response - I am not alone, so I hope this does not affect the results too much.'*
	'Non-clinical staff were asked to complete the survey but a third of the questions (those that asked about encounters with patients) were entirely irrelevant to this group as they do not encounter patients.'

6.2.6. Factor analysis and item reduction

The sample adequacy for PCA, as measured by the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, was excellent (Dziuban & Shirkey 1974) and the correlations between items, as measured by Bartlett's Test of Sphericity, were sufficiently large (<0.05) for a PCA (Dziuban & Shirkey 1974) (Table 57).

Table 57 – Tests of sample adequacy and sphericity of the field-test sample for principal component analysis

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.937
Bartlett's Test of Sphericity	Approx. Chi-Square	3991.053
	df	378
	Sig.	0.000

The three components extracted in the principal component analysis were accountable for more than 59% of the variance and the majority of the variance was explained by the first component (Table 58). As a rule of thumb, the component structure deemed to be acceptable if it cumulatively explains > 50% of the variance (Starkweather & Herrington 2011).

Table 58 – Variance explained by factors generated from the first field-test responses

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	12.807	45.738	45.738
2	2.238	7.994	53.733
3	1.591	5.683	59.415

All items clustered under the facets that they were intended to represent with factor loading coefficients of > 0.3. It was decided to psychometrically examine the top five items loaded in each component with the view of having 15 items in the final measure.

The five items loaded with the highest factor coefficients under each component demonstrated high internal consistencies as measured by Cronbach alpha (0.882 for workmanship, 0.901 for clinicianship and 0.784 for citizenship). In addition, the overall internal consistency of these 15 items was very high (Cronbach alpha 0.914). Therefore, the items within each facet and in the entire measure appeared to be highly correlated (Brown 2002). With five items in each facet and 15 items in the entire measure, all Cronbach alpha values are well-acceptable for a psychological measure of this kind (Brown 2002; Field 2009). Therefore, these 15 items were selected for the final version of the DBIP. (Table 59)

Table 59 – Loading of items under each of the three factors based on the field-test responses

(The items selected to represent each facet in the final measure are indicated in bold.)

	Factor 1 (Clinicianship)	Factor 2 (Workmanship)	Factor 3 (Citizenship)
Showing compassion towards patients	**0.778**		
Being attentive to the needs of patients	**0.775**		
Acting in a responsible fashion towards patients	**0.766**		
Communicating with patients in a clear and effective manner	**0.764**		
Treating patients fairly and without prejudice	**0.746**		
Being empathetic when caring for patients	0.741		
Respecting patients' confidentiality and privacy	0.720		
Providing advice to patients when required	0.715		
Respecting patients' autonomy	0.694	0.365	
Behaving in a reliable and dependable way	0.678		0.355
Treating colleagues of the same profession fairly and without prejudice		**0.775**	
Being receptive to constructive criticism		**0.768**	
Working well as a member of a team		**0.762**	
Treating other healthcare professionals fairly and without prejudice		**0.750**	
Having a positive attitude towards professional development		**0.712**	
Being able to manage situations where there is a conflict of interest	0.324	0.667	
Respecting colleagues of the same profession		0.662	
Reflecting on your actions with a view to self-improvement		0.584	
Being aware of own limitations		0.581	0.304
Making effective use of the resources available		0.560	0.320
Avoiding substance or alcohol misuse			**0.693**
Functioning according to the law			**0.675**
Adhering to professional rules and regulations	0.397		**0.649**
Taking a dedicated approach to work	0.383		**0.580**
Being sound in judgment and in decision making	0.398		**0.569**
Behaving honestly and with integrity	0.303		0.534
Being accountable for one's actions	0.393		0.471
Not using professional status for personal gain		0.449	0.467

6.2.7. Generalisability study of the final version of the Dundee Barometer of Institutional Professionalism

The D study using Generalisability Theory was conducted considering all 212 sets of data for the three domains (facets), 15 items and three questions per item. The G coefficient of 15 items was 0.927, which is excellent. The G study also illustrated that the highest variance component (0.647) was explained by respondents, and not by the questions, items or domains which indicated the effectiveness of the measure (Shavelson *et al.* 1989) (Table 60 and Table 61).

Table 60 - Variance explained by difference components according to a generalisability analysis of the final 15 items

Effect	Degree of freedom	Variance component
P*	**211**	**0.647**
S*	2	0.132
D*	2	(0.0)
I:D	12	0.196
PS	422	0.392
PD	422	(0.0)
PI:D	2532	0.390
SD	4	(0.0)
SI:D	24	0.255
PSD	844	0.105
PSI:D	5064	1.565

(*P=persons; S=questions; D=Facets)

Table 61- D study findings to estimate the reliability of the 15-item version

No. of facets	No. of items per facet	No. of questions per item	No. of items per facet	Mean	G-coefficient
3	1	3	1	0.10	0.72
3	2	3	2	0.05	0.84
3	3	3	3	0.04	0.88
3	4	3	4	0.03	0.91
3	**5**	**3**	**5**	**0.02**	**0.93**
3	6	3	6	0.02	0.94
3	7	3	7	0.02	0.95
3	8	3	8	0.02	0.95
3	9	3	9	0.01	0.96
3	10	3	10	0.01	0.96

In general, the 15 items selected for the final DBIP demonstrated moderate (0.41 – 0.8) inter-item correlation within the respective facet and mild (0 - 0.4) correlation between facets (Table 62).

Table 62- Inter-item correlations of the final Dundee Barometer of Institutional Professionalism items

		Workmanship					Clinicianship					Citizenship				
		a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**l**	**m**	**n**	**o**
Workmanship	a. Having a positive attitude towards professional development	1.00	0.61	0.63	0.53	0.56	0.42	0.46	0.42	0.44	0.37	0.42	0.30	0.19	0.42	0.39
	b. Working well as a member of a team	0.61	1.00	0.56	0.62	0.63	0.42	0.40	0.33	0.39	0.42	0.33	0.37	0.12	0.31	0.37
	c. Being receptive to constructive criticism	0.63	0.56	1.00	0.56	0.61	0.43	0.36	0.43	0.48	0.34	0.28	0.24	0.19	0.35	0.32
	d. Treating other healthcare professionals fairly and without prejudice	0.53	0.62	0.56	1.00	0.74	0.46	0.33	0.34	0.43	0.47	0.36	0.33	0.16	0.34	0.38
	e. Treating colleagues of the same profession fairly and without prejudice	0.56	0.63	0.61	0.74	1.00	0.46	0.44	0.39	0.51	0.53	0.44	0.41	0.18	0.37	0.44
Clinicianship	f. Communicating with patients in a clear and effective manner	0.42	0.42	0.43	0.46	0.46	1.00	0.66	0.62	0.64	0.65	0.39	0.26	0.15	0.42	0.50
	g. Acting in a responsible fashion towards patients	0.46	0.40	0.36	0.33	0.44	0.66	1.00	0.65	0.64	0.66	0.44	0.29	0.16	0.44	0.41
	h. Being attentive to the needs of patients	0.42	0.33	0.43	0.34	0.39	0.62	0.65	1.00	0.70	0.60	0.44	0.29	0.15	0.47	0.47
	i. Showing compassion towards patients	0.44	0.39	0.48	0.43	0.51	0.64	0.64	0.70	1.00	0.69	0.52	0.28	0.13	0.53	0.51
	j. Treating patients fairly and without prejudice	0.37	0.42	0.34	0.47	0.53	0.65	0.66	0.60	0.69	1.00	0.44	0.33	0.19	0.39	0.47
Citizenship	k. Adhering to professional rules and regulations	0.42	0.33	0.28	0.36	0.44	0.39	0.44	0.44	0.52	0.44	1.00	0.49	0.28	0.66	0.51
	l. Functioning according to the law	0.30	0.37	0.24	0.33	0.41	0.26	0.29	0.29	0.28	0.33	0.49	1.00	0.34	0.36	0.40
	m. Avoiding substance or alcohol misuse	0.19	0.12	0.19	0.16	0.18	0.15	0.16	0.15	0.13	0.19	0.28	0.34	1.00	0.36	0.41
	n. Being sound in judgment and in decision making	0.42	0.31	0.35	0.34	0.37	0.42	0.44	0.47	0.53	0.39	0.66	0.36	0.36	1.00	0.50
	o. Taking a dedicated approach to work	0.39	0.37	0.32	0.38	0.44	0.50	0.41	0.47	0.51	0.47	0.51	0.40	0.41	0.50	1.00

The items included to represent each of the three facets of the final version of DBIP are presented in Table 63.

Table 63 - Items included in the final version of Dundee Barometer of Institutional Professionalism

Facet of professionalism	Measured behaviour	Items of measurement
Workmanship	Individual-oriented *verses* team-oriented behaviour	1. Treating colleagues of the same profession fairly and without prejudice 2. Being receptive to constructive criticism 3. Working well as a member of a team 4. Treating other healthcare professionals fairly and without prejudice 5. Having a positive attitude towards professional development
Clinicianship	Practitioner-centred *verses* patient-centred behaviour	1. Showing compassion towards patients 2. Being attentive to the needs of patients 3. Acting in a responsible fashion towards patients 4. Communicating with patients in a clear and effective manner 5. Treating patients fairly and without prejudice
Citizenship	Self-centred *verses* society-centred behaviour	1. Avoiding substance or alcohol misuse 2. Functioning according to the law 3. Adhering to professional rules and regulations 4. Taking a dedicated approach to work 5. Being sound in judgment and in decision making

6.3. Discussion

The DBIP was field-tested in the Dundee Medical School to analyse its psychometric properties and acceptability, and the responses were used to finalise the content and format of DBIP. The methodology employed in the field-test and its findings are discussed below.

6.3.1. Methodology used

As in this study, it is a common practice for developers of psychological measures to use a convenience and easily accessible sample for their first field-test to determine the initial psychometric properties and finalise the content (Clark & Watson 1995). This helps identify the strengths and particularly the deficiencies of the measure before exposing it to larger and rather unfamiliar populations (Clark & Watson 1995). The homogeneity of such samples tends to help the initial analysis, but may hinder the opportunity of evaluating the performance of the measure with heterogenic samples (Clark & Watson 1995). In practical terms, it is not ethical to use a measure in multiple centres without at least some kind of evidence for its utility. The developers of the measures of the professionalism culture, discussed in the literature review, conducted their first field-tests in their own institutions. The responses that they received were used to analyse the psychometric characteristics of the respective measures. Therefore, the use of Dundee Medical School for the first field test for the DBIP was a rational approach. Measures like the DREEM (Roff *et al.* 1997) and the Jefferson Scales (Hojat *et al.* 2001; Hojat *et al.* 2006) have been tested subsequently with more heterogenic populations, e.g. internationally

(Roff 2005; Wetzel *et al.* 2010). This may be the next step for the DBIP, but this is beyond the scope of the doctoral project.

Although it was initially planned to administer the DBIP to medical students across all five years, it was sent ultimately to the fourth and final year students on the advice of the Dean of Teaching. Although the Dundee students are exposed to clinical environments from the first year onwards they are not fully integrated into clinical environments until the fourth and final years. Furthermore, this advice may have reflected the notion that professionalism among medical undergraduates develops in stages and, therefore, junior students may not possess a complete understanding of the entirety of the concept (Hilton & Slotnick 2005). However, the exclusion of junior students may have reduced the number of respondents but not the response rate as response rates for online surveys among different year groups appear to be very similar (Roff *et al.* 2012). The lack of knowledge among junior medical students about the content and context described by the items may result in non-attentive responding and unreliable results (DeVellis 2003, pp.69-70). Therefore, the compliance to the advice of the Teaching Dean may have improved the quality of responses though the quantity may have been compromised to a certain degree.

6.3.2. Findings

The findings are discussed in relation to: the responses received; the internal structure of the measure; the items, facets and measure scores; the differences observed between demographic groups; the feedback of the respondents; and the outcome of the item reduction process.

6.3.2.1. The response rate and the number of responses

a. Response rate

The response rate is important to interpret the professionalism culture of the institution (Baldwin & Daugherty 2006). The response rate among the students (25%) and the staff (19%) of the Dundee Medical School for the DBIP was rather low. The low response rates among the staff and students may be attributable to multiple factors. Although there are occasional findings to the contrary (Braithwaite *et al.* 2003), the response rates for web-based questionnaires are generally lower than the paper-based postal questionnaires especially among medical professionals (Shih & Fan 2008) and students (Paolo *et al.* 2000). The measures of professionalism culture, delivered by both postal (Hojat *et al.* 2009a) and face-to-face (Kalet & Steven 2004; Quaintance *et al.* 2008) modes, achieved relatively high response rates. This, however, may not necessarily be due to the preference of respondents towards paper-based surveys; a web-based format is considered more attractive and interesting compared to paper formats (Braithwaite *et al.* 2003). The respondents (staff and students of educational institutions) may have developed questionnaire fatigue; they are frequently bombarded with surveys of varying levels of quality for various purposes (Baruch & Holtom 2008; Porter *et al.* 2004). The fast-paced

educational culture and growing demands on time may have also made the staff and students less willing to commit to completing a survey (Baruch & Holtom 2008). In addition to external factors, the length of the DBIP, which required flipping through 32 screens to complete, may have adversely affected the response rate, of which survey length is a clear determinant (Sax *et al.* 2003). The length was a 'matter of concern' highlighted in respondents' evaluation of the DBIP which is discussed below under evaluation findings.

The response rate of medical students for the DBIP (25%), however, was in par with similar studies. Roff *et al* (2012), who conducted a web-based Poly-professionalism inventory in the same institution in 2009, were able to achieve a very similar response rate (22%) for fourth year students (final year students were not included in the study). It appears that, in UK medical school settings, low response rates among senior medical students to surveys, especially to web-based surveys, are not uncommon (Paolo *et al.* 2000). However, Wiggleton *et al* (2010) achieved 60% response rate among the fourth-year students of the Vanderbilt Medical School, USA for their web-based moral distress survey. Unlike the DBIP or the Poly-professionalism inventory (Roff *et al.* 2012), Wiggleton *et al* deployed an increased involvement of students in the development process and generous incentives to respondents by means of a lottery draw with the view of improving response rate. Apart from possible cultural differences between the two study settings (UK and USA) (Cook *et al.* 2009), the high response rate observed in the moral distress survey may well be attributed to the involvement of students in the development process,

as incentives demonstrated no significant impact on response rates in academic settings (Baruch & Holtom 2008; Sax *et al.* 2003).

The response rates of the UK healthcare professionals to web-based surveys in general may range from nine to 94% (Braithwaite *et al.* 2003). Universally, if these surveys are carried out in an institutional environment, e.g. a medical school, the expected response rate would be around 35.7% (SD +/- 18.8) (Baruch & Holtom 2008). The rate of responses among staff members (19%) to the DBIP, therefore, was observed to be on the lower end of the range. With the academic staff of the Dundee Medical School, Roff *et al* (2012) achieved a response rate of 22%. The number of staff members responded to the DBIP (132), however, was much higher than to the survey (57) of Roff *et al* (2012).

Despite the deployment of strategies to enhance responding the response rate to DBIP was low. This impacted on the interpretation and generalisation of the total and domain scores and the comparison of demographic groups. It may be a drawback of the web-based approach to the DBIP. The alternative method of ensuring high response rate, at least among students, is delivering the paper-based version face-to-face after an educational activity, e.g. after a lecture or examination. However, in paper-based format, it is not practical to incorporate many features of the DBIP which enhance the quality, efficiency and effectiveness. In addition, the options available for designing, delivery and analysis are comparatively limited with a paper version. Low response rates may be causing non-response bias in results, i.e. there may be a possibility of underrepresenting the perceptions of non-respondents and overrepresenting the perception of respondents

(Sax *et al.* 2003). However, a meta-analysis on survey research by Cook *et al* (2000) demonstrated that the representativeness of the respondents has more impact than the response rate itself on the credibility of findings. Therefore, in future applications of the DBIP, where the primary emphasis is the exploration and interpretation of the institutional professionalism culture, low response rates should be anticipated and the use of a representative sampling technique, e.g. quota sampling, should be recommended. In addition, the DBIP can be made a formal component of professionalism education of the institution which may enhance the response rates. However, it should not be forced unnecessarily but should be used judiciously as respondents may develop apathy and negative attitude towards the DBIP.

As the exact demographic details of staff and student population required in this study could not be accessed, an analysis of results in terms of institutional culture was not carried out.

b. Number of responses

In the first field test, as the emphasis was on establishing the internal structure of the measure, the number of responses was more important than the response rate (Clark & Watson 1995). Although there is no empirically determined minimum number of responses for conducting factor and psychometric analyses of measures, rules of thumb have been proposed based on the theoretical underpinning of the respective analysis (van Voorhis & Morgan 2007). With regard to factor analyses, more than 300 responses are considered ‘good’ and between 200 and 300 are considered ‘acceptable’ (Clark & Watson

1995; DeVellis 2003, pp.137; van Voorhis & Morgan 2007). In fact, it has been suggested by the experts that 200 responses are adequate for a meaningful factor analysis if the number of items are less than 40 (Clark & Watson 1995; DeVellis 2003, pp.137; van Voorhis & Morgan 2007). For determining relationships, e.g. correlations, around 50 responses are sufficient (van Voorhis & Morgan 2007). The 28-item DBIP was completed by 212 which was sufficient to perform factor analysis and, in some instances, demographic comparisons.

6.3.2.2. Internal structure and statistical characteristics of the measure

The number and the quality of items primarily and interactively determine the internal consistency of a measure; for example, a large number of poorly written items or a few well-written items may produce high internal consistency (Brown 1997). The latter appears to be the most appropriate explanation for the high internal consistency found in this study as the number of items included in the measure was not more than 30 (Brown 1997), the high internal consistency should indicate good quality of items and a positive correlation between them.

The distribution of the total scores did not show a normal distribution, but was highly skewed towards the higher end of the scale. It is not uncommon for assessment or measure scores to be skewed due to existence or non-existence of measured characteristics (Field 2009, pp.540-551). Therefore, this may reflect either the actual existence of a very healthy institutional professionalism culture in relation to all facets

measured. Conversely, it may also be due to socially desirable responding. This issue is discussed in the next section.

6.3.2.3. Item, domains and total scores of the measure

The developers of DREEM introduced a template to help interpret its scores hence the educational environment (0-0-50 very Poor, 51-100 plenty of problems, 101-150 more positive than negative, 151-200 excellent) (Roff *et al.* 1997). Many other measures of professional culture, however, do not have such templates or guidelines to interpret the scores (Arnold *et al.* 1998; Hojat *et al.* 2001; Quaintance *et al.* 2008; Wiggleton *et al.* 2010). Although such templates may be helpful, it is totally fictional to include such a template for the DBIP after the first-field test. However, it may be considered in the future after the application of the latest version of the DBIP and analysis of the response patterns.

In the first field-test, the scores of every item tended to be towards the higher end of the continuum contributing to very high scores for the three facets and for the entire measure (Table 51, Figure 12 and Table 51). Although the response rate was inadequate to make conclusive remarks on the overall cultural outlook, which is a limitation of the first field test, it appears that the professionalism culture of the Dundee Medical School is very healthy and beneficial to all stakeholders, i.e. the respondents indicate that the professionalism culture of Dundee Medical School is more society centred than self-centred, more patient-centred than a practitioner-centred, and more team-work oriented than an individual-oriented. In relative terms, the total scores suggest that the social

culture (citizenship) is the healthiest aspect in the institutional professionalism. The members of the institution indicated a strong intention, thus a high probability, of acting with accountability, probity, dedication and law-abiding manner, and making sound judgements. The working culture (workmanship), on the other hand, showed the lowest total score. It is apparent in the analysis of individual items scores that the members of the institution hold relatively weaker intentions on managing conflicts of interests, being open to constrictive criticisms, engaging in professional development activities, making effective use of resources, being a reflective practitioner and making use of the 'privileged' position of a doctor for personal benefits. All three aspects (personal attitude, institutional expectation and achievability) with regards to these behaviours were scored relatively low; the achievability scores were the lowest. The total score for the culture of patient-centeredness lies in between the social and working culture. In fact, there is no much difference between the total scores for the social and clinical cultures. There was a noticeable general pattern in the attitude, expectation and achievability scores across the three facets. The highest and lowest were the personal attitude and the achievability scores respectively, and the expectation scores remained in between. The respondents were of the view that although they value behaving professionally, the institutional expectation towards such behaviours and the conduciveness for behaving such ways in the institutional environment were somewhat weak. According the TPB, this may undermine the materialisation of professionalism behaviours in the institutional environment (Ajzen 1991). The reasons may be the lapses in emphasising what is expected, the workload or the time pressures. As the challenges faced by the respondents

were not clearly apparent with the DBIP scores, it may be necessary to explore such issues with qualitative methods such as interviews or focus group discussions. It appears that the DBIP is effective at identifying even subtle differences in various aspects of the professionalism culture of a given institute. These aspects can be focused and emphasised if interventions are to be planned and implemented.

Such high scores for items and facets in a survey, however, could have also been produced by the ceiling effect mainly due to socially desirable responding which is common with professionalism attitude surveys (Baldwin & Daugherty 2006). As discussed in Chapter 5, the use of TPB as the basis for the development of the rating scale should have minimised the effect of socially desirable responding. In this study, the scores for each TPB component provided evidence to support this; the scores for the three components were consistently different from each other. Mostly the 'attitude' scores were the highest and the 'achievability' scores were the lowest and 'expectation' scores remained in between (Table 50 & Figure 11). The responses to the attitude question may have been affected by social desirability as questions on self are more prone to socially desirable responses than others (Baldwin & Daugherty 2006). It is unlikely that the other two questions were affected by social desirability. The possible bias caused by the response to the attitude question may have been countered, at least to certain extent, by the other two questions. Although this suggests that the use of a theory-based approach minimised certain ill effects commonly associated with psychological measures, future application of the DBIP in multiple centres will be necessary to produce conclusive evidence. The ability of the DBIP to discriminate between institutions would provide such evidence.

The correlations between facets were statistically significant and moderate (Table 52). The moderate correlations between facets indicate the actual existence of sub-domains within the construct (professionalism culture in this instance) (Clark & Watson 1995). If the correlations were too low or negative the sub-domains themselves may be independent constructs, and if they were too high they may not be subdomains but all items may represent the construct itself; the division of the construct into subdomains would become artificial and meaningless (Clark & Watson 1995). The facets, therefore, actually represent the entirety of professionalism culture.

6.3.2.4. Comparison of demographic groups

Although the item scores of students were marginally higher than the faculty, the scores of both groups for all items were on the high end. The picture portrayed by the results suggests that the students and the faculty of the Dundee Medical School have very positive attitudes towards professionalism behaviour (Roff *et al.* 2011). In this study, the students perceived the professionalism of the Dundee Medical School strongly with significant positivity compared to the faculty. However, Roff *et al* (2011) observed a mixed pattern in the same institution; students and faculty perceived the strength of the various aspects of professionalism differently. Quaintance *et al (2008)* observed significantly higher scores from students than faculty for their Professionalism Climate Instrument. In the present study it was observed that female faculty have strongly positive attitudes towards the relationship of doctors with patients and co-workers compared to their male counterparts. Similar observations have been with regard to moral distress (Wiggleton *et al.* 2010) and empathy (Hojat *et al.* 2003a).

The ability to compare demographic sub-groups was restricted due to the low response rate, which can be considered a significant limitation of this field test. Certain publications, which reported the development and validation of measures on professionalism culture compared, analysed and discussed the prevailing situation among the respondents based on their first field-test findings (Hojat *et al.* 2009a; Quaintance *et al.* 2008; Roff *et al.* 2012; Wiggleton *et al.* 2010). Others, however, irrespective of response rates, primarily focused on determining the psychometric properties of the respective measures with their first-field test responses (Arnold *et al.* 1998; Blackall *et al.* 2007; Hojat *et al.* 2001; Roff *et al.* 1997; Thrush *et al.* 2011). In the current study, with a low response rate, it was inappropriate to analyse responses in order to make general comments on the similarities and differences between demographic groups. However, with the limited comparisons conducted, it is apparent that the DBIP has the capability of identifying apparent as well as subtle differences between groups.

6.3.2.5. Feedback of respondents towards the measure

The reaction of respondents towards psychological measures is rarely reported in the literature. However, evaluating the reaction of respondents to any measure is always advised and these reactions could provide important information on the acceptability of the measure (van der Vleuten 1996). The evaluation suggests that the DBIP was largely acceptable to the targeted respondents. The majority perceived that there was: a similarity between the professionalism attributes included in the DBIP and their own understanding of professionalism; a representativeness of the overall picture of the professionalism culture generated from the responses of individuals; and an

appropriateness in the format and presentation. All these aspects endorse the face validity of the measure. The length of the measure, however, appeared to be a cause for concern. Lengthy measures may reduce the credibility of responses (DeVellis 2003 pp.96-100). This may have impacted on the overall acceptability of the measure, which was an inconclusive aspect of the evaluation. Although they were few in numbers, the open comments reflected several important issues; the confusing nature of the term 'as a part of wider society' at least for some respondents, and the DBIP should not be sent to non-clinical staff as they have limited exposure to / knowledge of clinical environments. The comments on the DBIP content were not specific enough for changes to be made.

6.3.2.6. Factor analysis and item reduction

The principal component analysis with the specification to extract three factors was a rational approach to establish the validity of the initial component structure (Brown 2009a). The data demonstrated the necessary statistical rigour for such analysis (Field 2009, pp.627-685). The similarity between both principal component analyses (public survey and first field-test) provided supportive evidence for the construct validity of the measure (Clark & Watson 1995). However, a confirmatory factor analysis, using a large number of respondents in various settings, will be necessary in the future to provide robust evidence for construct validity (Clark & Watson 1995).

The items loaded with high coefficients {>0.5(Brown 2009a)} under their respective domains. The information generated by principal component analysis is useful in making informed decisions, especially in reducing the number of items (Brown 2009b). The

number of items should strike a balance between the reliability, validity and practicality of a measure (DeVellis 2003, pp.98-100). Based on empirical evidence it was decided to select 15 items basing on the PCA. The decision study conducted using the generalisability theory predicted that a very high reliability could be achieved with five items selected for each facet. On subjective judgement, it appears that these items represent the facets they belong to and professionalism in general; therefore, the validity should not be compromised. From the perspective of respondents, expecting 45 responses (15 items with three questions to each) is within the range of item numbers included in the final versions of the professionalism culture measures, e.g. DREEM (50 items), Professionalism Climate Instruments (32 items), Poly-Professionalism Inventory I (30 items), Moral Distress Survey (50 items). In terms of number of items, DBIP (15) is very similar to the ABIM Professionalism Scale (12), and the Jefferson Scales of Lifelong Learning Scale (14) and the Jefferson Scale of Empathy (20). Fifteen items demonstrate excellent reliabilities at domain and measure levels without compromising the content validity which should be the primary psychometric goal in finalising the number of items in a measure (Clark & Watson 1995; DeVellis 2003, pp.98-99). The moderate to high correlations between intra-facet items and the mild correlation between inter-facet items indicate the presence of a sound internal structure within the measure, i.e. confirm the meaningfulness of high reliabilities (Clark & Watson 1995). Therefore, it was appropriate to reduce the number of items to 15.

6.4. Conclusion

The final stage of the study helped develop the measure of the professionalism culture, named as the Dundee Barometer of Institutional Professionalism (DBIP), based on the findings of the previous components of this study. The subsequent field test demonstrated sound psychometric properties and acceptability for the DBIP, which was subsequently reduced to 15 items; five items representing each of the three facets. Each item has three questions. Both the statistical analysis and evaluation findings suggested that, except for the interpretation of citizenship, 'as a part of wider society', the remainder of facet names, their interpretations, the wording of items and questions were appropriate. These were kept unchanged. The evaluation, however, indicated that the facet of citizenship needed to be interpreted differently. Therefore, the phrase 'as a part of wider society', which provided context to items under the facet of citizenship, will be replaced with 'as a member of your profession' in the future version of the DBIP. In order to enhance the credibility of responses, the DBIP will be recommended for use with a representative sample of medical undergraduates with adequate exposure to clinical environments and the faculty, who are involved in clinical practice, of a given institution.

Section D – Overall Conclusions

This section discusses the overall observations on the process (the development of the DBIP) and the product (the DBIP as a measure of professionalism culture) of this doctoral project.

Chapter 7 – Strengths, uses, limitations of the Dundee Barometer of Institutional Professionalism and areas of future research

In this doctoral project, a quantitative measure of institutional professionalism culture (Dundee Barometer of Institutional Professionalism) was developed and validated for use in medical schools in the UK. The validation involved a thorough literature review and achieving consensus among the public and medical professionals. The measure demonstrated high reliability and acceptability in its first-field test which helped finalise the 15-item DBIP. The key strengths, uses, limitations, and future research associated with DBIP are summarised in this chapter.

7.1. Strengths

The DBIP is the first qualitative measure of institutional professionalism culture developed with national consensus of the UK public and medical practitioners, and a sound theoretical basis, which underpins the concept of institutional professionalism culture. Therefore, it addresses two important concerns repeatedly emerging in the professionalism literature; not grounding professionalism measurements on the socio-cultural context, and the lack of theoretical underpinning in the existing survey instruments questioning their credibility (Archer *et al.* 2008; Jha *et al.* 2007; Rees & Knight

2007). These features of the DBIP also fulfilled the primary requirements of establishing the content validity of measures (DeVellis 2003, pp.49&50).

In the process of developing the DBIP, a comprehensive and meaningful framework to conceptualise the construct of professionalism culture (the three facet model) emerged. Therefore, the DBIP made the abstract concept of professionalism culture meaningful and measurable (DeVellis 2003). In this aspect, the DBIP is similar to the Jefferson scales which defined the concepts of empathy and lifelong learning in measurable terms (Hojat *et al.* 2002; Hojat *et al.* 2003b).

As observed in the evaluation, the content, format and presentation of the DBIP was acceptable to most of the respondents. The revised version of DBIP will hopefully make the DBIP popular not only within the UK but also internationally.

The DBIP complies very strongly with the current trend for delivering assessments in computer/web-based modes. The online mode gives the respondents the flexibility and adaptability which may be needed in time-pressured fields like medicine. It gathers analyses and interprets responses of large groups of students and staff within a short period of time. This feature is particularly important in the current context where a need for such cost-effective measurements of professionalism is frequently expressed (Roff *et al.* 2009; van Mook *et al.* 2009a).

7.2. Uses

DBIP can be used as a tool for screening, comparing, evaluating and understanding the institutional professionalism culture. More than one of these purposes could be achieved

with a single administration. A profile of institutional professionalism culture of a particular medical school can be screened cost-effectively to detect its positive and negative aspects. The differences between different clinical contexts, cohorts, demographic groups in the same medical school / programme can be compared. The interventions can be planned accordingly and the impact of such interventions can be evaluated by the re-administration of the DBIP. The findings can be used as a performance indicator of the overall quality of the educational programme of the institution. The impact or the relationship between professionalism culture and student performance, attrition rates or stress levels can also be measured using the DBIP. As there are validated measures for some of these aspects (e.g. stress levels), the scores of those measures and the DBIP can be conveniently compared.

It may also be useful for comparing medical schools, e.g. schools with different curricula designs, different approaches to professionalism teaching / learning, school entrants and graduate entrants. When the DBIP is validated for use outside the UK, it will help understand cultural differences in professionalism culture.

As for the DREEM (Roff 2005) and the Jefferson Sale of Empathy (Hojat *et al.* 2003a), the DBIP scores of individual respondents may be compared with their academic ability, learning styles or emotional states (e.g. stress) as measured by validate instruments. The conflict between protection of anonymity and obtaining genuine perceptions should be carefully managed when the DBIP is used for such purposes.

In every application of the DBIP, the sensitivities of the respondents, institutions and authorities should be carefully considered before sharing them with wider audiences. The respondents may not reveal the truth but safeguard the reputation of their institution by providing socially desirable responses if the trust between the DBIP administers and potential respondents are not maintained.

7.3. Limitations

The limitations pertinent to each stage of this project were discussed in depth under the relevant sections. The key elements of those limitations are summarized here before discussing the potential limitations of the DBIP in general.

Summary of the limitations pertinent to the development process

An international group of medical educationalists were involved in the refinement of items, which were included in the public and professionals' surveys, for the UK context. This may have enhanced the potential of using the DBIP internationally. However, as the interpretation of professionalism is influenced by the cultural background, the process may have been affected by the cultural mix. In addition, as the author trained outside the UK his interpretation of items may have influenced by his cultural background.

In the validation process, the use of consensus survey method was helpful to obtain the perceptions of wider, larger and more representative groups of the public and professionals. However, if the Delphi technique was used, an iterative dialogue between the participants throughout the prioritisation process would have been maintained which may enhance the credibility of the final list of items.

In the public survey, the use of a selected group signed up to participate in research and surveys improved the accessibility of a representative sample and improved the credibility of responses. However, the literacy level of this group may be higher than the ordinary public and this may have caused bias in responses. Similarly, the use of a selected group of medical educationalists to represent 'medical professionals' may not be representative of the profession in general.

In the development of the DBIP, the recommendations for the development of quantitative psychological measures were followed almost to the very letter. However, the validation process could have been improved by incorporating a qualitative component. For example, an analysis of professionalism dilemmas expressed by students in individual or focus group interviews could have been an alternative method of identifying the attributes. Such qualitative approach may improve the depth of the validation process.

The interpretation of the total and domain scores for the Dundee Medical School and the comparison of its demographic groups were restricted by the low response rates observed in the first field test.

Limitation of the DBIP as a measure of professionalism culture

The DBIP may be vulnerable to the weaknesses of perception-based surveys to some extent. The major concerns about measures like the DBIP include: the difference which may exist between perceptions and the real situation; the influence of social desirability on responses; non-response bias; low response rate; and difficulty with causal

interpretations (Baldwin & Daugherty 2006). As discussed in Section 3, it can be anticipated that the use of a theory-based rating scale enhanced the credibility and social desirability of responses. However, conclusions on the credibility of response can only be made after comparing the findings of the DBIP and an observation study in the same institution, conducted in parallel. The low response rate and non-response bias can be a continuing issue with the use of web-based format. Targeting a representative sample of the population concerned (e.g. use of a quota sampling technique) may help minimising the non-response bias caused by low response rates (Cook *et al.* 2000). The difficulty of making causal relationship, i.e. the problems but not the reasons can be identified, is an inherent problem with survey based methods (Baldwin & Daugherty 2006). This limitation may be addressed with follow up qualitative approaches, e.g. focus group discussions, to further explore the areas of concern (Cohen & Manion 1994, pp.284-286). However, using the DBIP, which is less resource-intensive, as a screening tool initially and investing resources on more resource-intensive qualitative approaches only if problems exist may be an effective and efficient strategy. Alternatively, the DBIP can be improved to provide opportunities for respondents to express their concerns on professionalism culture as qualitative comments. This will also enhance the authenticity of responses and help obtain more comprehensive picture of the professionalism culture of an institution.

It may be argued that the web-based mode limits the wide use of the measures due to accessibility and literacy issues. However, this argument is invalid in the UK context given the wide access to computers and internet with very high computer literacy rate (Seago *et*

al. 2002). In an event, where the DBIP is validated for use in a developing country, this argument should be duly considered (Ameh *et al.* 2008).

7.4. Future research

There are several future steps required for the further development of the DBIP. It needs to be used widely in UK medical schools to establish its ability to discriminate between medical schools. Publication of preliminary results will help attract authorities in other medical schools in the UK to apply this in their respective medical schools. In addition, the authorities of the 32 medical schools in the UK, e.g. Deans, will be approached with written invitations requesting their permission and support to administer the DBIP in their respective medical school.

The pooled findings of a wider use can be utilised to conduct a confirmatory factor analysis, which has potential to provide strong evidence for the construct validity of the DBIP as a measure of professionalism culture. It can be further validated for use in other healthcare disciplines such as dentistry and nursing at undergraduate level in the UK. The validity of the DBIP can also be extended to postgraduate contexts of healthcare disciplines.

In parallel, it can be validated for use internationally, similar to Jefferson Scales (Hojat *et al.* 2003a; Hojat *et al.* 2009a) and DREEM (Roff 2005) to make it a generic measure of institutional professionalism culture. However, the validity of the specific content of the DBIP in respective countries needs to be determined as it is UK-based at present. The cultural differences may affect the credibility and acceptability of the DBIP. Therefore, the

conceptualisation underpinning the measure should be evaluated thoroughly in relation to each cultural context. However, as observed in the international survey conducted as a part of the validation process, it was observed that there is a set of core attributes of professionalism valued by all cultural groups around the world (Chandratilake *et al.* 2012). Therefore, the DBIP, as a concept, will be acceptable to many countries. The initial steps can be taken after publishing the initial findings in a reputed international journal. The presentations made on the DBIP in international conferences have already stimulated several international collaborators from countries like Australia, Thailand and Indonesia.

References

ABIM (2000). *Professionalism in Medicine: Issues and Opportunities in the Educational Environment.* Philadelphia: American Board of Internal Medicine. 5 - 10.

ABIM (2001). *Project Professionalism.* Philadelphia: American Board of Internal Medicine. 1 - 31.

Ainsworth MA & Szauter KM (2006). Medical student professionalism: Are we measuring the right behaviors? A comparison of professional lapses by students and physicians. *Academic Medicine, 81*(10 Suppl), s83-86.

Ajzen I (1991). The theory of planned behaviour. *Organisational Behaviour and Human Decision Processes, 50*, 179-211.

Ajzen I (2002). Constructing a TpB questionnaire: conceptual and methodological considerations. Retrieved 12.12.2010, from http://www-unix.oit.umass.edu/~aizen/pdf/tpb.measurement.pdf.

Ameh N, Kene TS & Ameh EA (2008). Computer knowledge amongst clinical year medical students in a resource poor setting. *African Health Sciences, 8*(1), 40-43.

Andre J, Brody H, Fleck L, Thomason CL & Tomlinson T (2003). Ethics, professionalism, and humanities at Michigan State University College of Human Medicine. *Academic Medicine, 78*(10), 968-972.

Ang M (2002). Advanced communication skills: conflict management and persuasion *Academic Medicine, 77*(11), 1166.

Aramesh K, Mohebbi M, Jessri M & Sanagou M (2009). Measuring professionalism in residency training programs in Iran. *Medical Teacher, 31*(8), e356-361.

Archer R, Elder W, Hustedde C, Milam A & Joyce J (2008). The theory of planned behaviour in medical education: a model for integrating professionalism training *Medical Education, 42*(8), 771 - 777.

Armitage CJ & Conner M (1999). Predictive validity of the theory of planned behaviour: The role of questionnaire format and social desirability. *Journal of Community & Applied Social Psychology, 9*(4), 261-272.

Armitage CJ & Conner M (2001). Efficacy of theory of planned behaviour: a meta analytic review *British Journal of Social Psychology, 40*(4), 471 - 499.

Arnold E, Blank L, Race K & Cipparrone N (1998). Can professionalism be measured? The development of a scale for use in the medical environment. *Academic Medicine, 73*(10), 1119-1121.

Arnold L (2002). Assessing professional behaviour: yesterday, today, and tomorrow. *Academic Medicine, 77*(6), 502-515.

Arnold L, Shue CK, Kalishman S, Prislin M, Pohl C, Pohl H *et al.* (2007). Can there be a single system for peer assessment of professionalism among medical students? A multi-institutional study. *Academic Medicine, 82*(6), 578-586.

Arnold L & Stern DT (2006). What is medical professionalism? In Stern DT (Ed.), *Measuring Medical Professionalism* (pp. 15-37). New York: Oxford University Press.

Arnold RM & Forrow L (1993). Assessing competence in clinical ethics: are we measuring the right behaviors. *Journal of General Internal Medicine, 8*(1), 52-54.

Austin EJ, Evans P, Magnus B & O'Hanlon K (2007). A preliminary study of empathy, emotional intelligence and examination performance in MBChB students. *Medical Education, 41*(7), 684-689.

Baernstein A, Oelschlager AM, Chang TA & Wenrich MD (2009). Learning professionalism: perspectives of preclinical medical students. *Academic Medicine, 84*(5), 574-581.

Baldwin DC & Daugherty SR (2006). Using surveys to assess professionalism in individuals and institutions In Stern DT (Ed.), *Measuring Medical Professionalism* (pp. 95 - 116). New York: Oxford University Press.

Baldwin DC & Self DJ (2006). The assessment of moral reasoning and professionalism in medical education and practice In Stern A (Ed.), *Measuring Medical Professionalism* (pp. 75-94). New York: Oxford University Press.

Bandura A (1989). Human agency in social cognitive theory. *American Psychologist, 44*(9), 1175-1184.

Bardecki MJ (1984). Participants response to the Delphi Method an attitudinal perspective. *Technological Forecasting and Social Change, 25*(3), 281-292.

Baruch Y & Holtom BC (2008). Survey response rate levels and trends in organizational research. *Human Relations, 61*(8), 1139-1160.

Beaudoin C, Maheux B, Cote L, Des Marchais JE, Jean P & Berkson L (1998). Clinical teachers as humanistic caregivers and educators: perceptions of senior clerks and second-year residents. *Canadian Medical Association Journal, 159*(7), 765-769.

Benbassat J & Baumal R (2004). What is empathy, and how can it be promoted during clinical clerkships? *Academic Medicine, 79*(9), 832-839.

Ber R & Alroy G (2002). Teaching professionalism with the aid of trigger films. *Medical Teacher, 24*(5), 528-531.

Bhugra D (2008). Professionalism and psychiatry: the profession speaks. *Acta Psychiatrica Scandinavica, 118*(4), 327-329.

Black J, Girolami S, Woodbury G, Hill M, Contreras-Ruiz J, Whitney JD *et al.* (2011). Understanding pressure ulcer research and sducation needs: a comparison of the association for the advancement of wound care pressure ulcer guideline evidence levels and content validity scores. *Ostomy Wound Management, 57*(11), 22-35.

Blackall GF, Melnick SA, Shoop GH, George J, Lerner SM, Wilson PK *et al.* (2007). Professionalism in medical education: the development and validation of a survey instrument to assess attitudes toward professionalism. *Medical Teacher, 29*(2-3), e58-62.

Blue AV, Crandall S, Nowacek G, Luecht R, Chauvin S & Swick H (2009). Assessment of matriculating medical students' knowledge and attitudes towards professionalism. *Medical Teacher, 31*(10), 928-932.

Bockers A, Jerg-Bretzke L, Lamp C, Brinkmann A, Traue HC & Bockers TM (2010). The gross anatomy course: an analysis of its importance. *Anatomical Sciences Education, 3*(1), 3-11.

Boelen C & Woollard B (2009). Social accountability and accreditation: a new frontier for educational institutions. *Medical Education, 43*(9), 887-894.

Boenink AD, de Jonge P, Smal K, Oderwald A & van Tilburg W (2005). The effects of teaching medical professionalism by means of vignettes: an exploratory study. *Medical Teacher, 27*(5), 429-432.

Bombeke K, Symons L, Debaene L, De Winter B, Schol S & van Royen P (2010). Help, I'm losing patient-centredness! Experiences of medical students and their teachers. *Medical Education, 44*(7), 662-673.

Bone A, Ncube F, Nichols T & Noah ND (2008). Body piercing in England: a survey of piercing at sites other than earlobe. *British Medical Journal, 336*(7658), 1426-1428.

Boon K & Turner J (2004). Ethical and professional conduct of medical students: review of current assessment measures and controversies. *Journal of Medical Ethics, 30*(2), 221-226.

Borgstrom E, Cohn S & Barclay S (2010). Medical professionalism: conflicting values for tomorrow's doctors. *Journal of General Internal Medicine, 25*(12), 1330-1336.

Braithwaite D, Emery J, de Lusignan S & Sutton S (2003). Using the Internet to conduct surveys of health professionals: a valid alternative? *Family Practice, 20*(5), 545-551.

Branch WT (2005). Use of critical incident reports in medical education - A perspective. *Journal of General Internal Medicine, 20*(11), 1063-1067.

Branch WT, Jr. (2000). Supporting the moral development of medical students. *Journal of General Internal Medicine, 15*(7), 503-508.

Branch WT, Jr. (2010). The road to professionalism: reflective practice and reflective learning. *Patient Education And Counseling, 80*(3), 327-332.

Brazeau CM, Schroeder R, Rovi S & Boyd L (2010). Relationships between medical student burnout, empathy, and professionalism climate. *Academic Medicine, 85*(10 Suppl), s33-36.

Brennan T, Blank L, Cohen J, Kimball H, Smelser N, Copeland R *et al.* (2002). Medical professionalism in the new millennium: a physicians' charter. *Lancet, 359*(9305), 520-522.

Brockbank S, David TJ & Patel L (2011). Unprofessional behaviour in medical students: A questionnaire-based pilot study comparing perceptions of the public with medical students and doctors. *Medical Teacher, 33*(9), e501-508.

Brown JD (1997). Reliability of surveys. *JALT Testing and Evaluation SIG Newsletter, 1*(2), 18-21.

Brown JD (2002). The Cronbach alpha reliability estimate. *JALT Testing & Evaluation SIG Newsletter, 6*(1), 17 - 18.

Brown JD (2009a). Choosing the right type of roation in PCA and EFA. *JALT Testing and Evaluation SIG Newsletter, 13*(3), 20 - 25.

Brown JD (2009b). Principal component analysis and exploratory factor analysis: definitions, differences and choices *JALT Testing and Evaluation SIG Newsletter, 13*(1), 26-30.

Brownell AK & Cote L (2001). Senior residents' views on the meaning of professionalism and how they learn about it. *Academic Medicine, 76*(7), 734-737.

Bryan CS & Babelay AM (2009). Building character: a model for reflective practice. *Academic Medicine, 84*(9), 1283-1288.

Bryan RE, Krych AJ, Carmichael SW, Viggiano TR & Pawlina W (2005). Assessing professionalism in early medical education: experience with peer evaluation and self-

evaluation in the gross anatomy course. *Annals Academy of Medicine, Singapore, 34*(8), 486-491.

Buyx AM, Maxwell B & Schone-Seifert B (2008). Challenges of educating for medical professionalism: who should step up to the line? *Medical Education, 42*(8), 758-764.

Carifio J & Perla R (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42*(12), 1150-1152.

Casper ES (2008). Using implementation intentions to teach practitioners: changing practice behaviour via continuing education *Psychiatric Services, 59*(7), 747 - 752.

Castillo JJ (2009). Quota sampling. Retrieved 10.09.2009, from http://www.experiment-resources.com/quota-sampling.html.

Chandratilake M, McAleer S & Gibson J (2011). *Development and validation of an inventory to measure professional culture of institutions training healthcare undergraduates: Dundee Barometer of Institutional Professionalism (DBIP).* Paper presented at the Annual Scientific Meeting of the Association for the Study of Medical Education Royal College of Physicians, 13 – 15 July, Edinburgh, UK.

Chandratilake M, McAleer S & Gibson J (2012). Cultural similarities and differences in medical professionalism: a multi-region study. *Medical Education, 46*(3), 257-266.

Chandratilake M, Roff S, McAleer J & Gibson J (2010). *Inventory I: academic integrity for junior undergraduates in the health professions: development phase of Dundee PolyProfessionalism Inventory.* Paper presented at the Annual Conference of Association for Studies Medical Education (ASME), The Royal college of Physicians, 15 – 17 July,Edinburgh, UK.

Chard D, Elsharkawy A & Newbery N (2006). Medical professionalism: the trainees' views. *Clinical Medicine, 6*(1), 68-71.

Chretien KC, Greysen SR, Chretien JP & Kind T (2009). Online posting of unprofessional content by medical students. *Journal of American Medical Association, 302*(12), 1309-1315.

Christner JG, Stansfield RB, Schiller JH, Madenci A, Keefer PM & Pituch K (2010). Use of simulated electronic mail (e-mail) to assess medical student knowledge, professionalism, and communication skills. *Academic Medicine, 85*(10 Suppl), S1-4.

Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P *et al.* (2006). Rating scales, scales of measurement, issues of reliability - Resolving some critical issues for clinicians and researchers. *Journal of Nervous and Mental Disease, 194*(8), 557-564.

Clark LA & Watson D (1995). Constructing validity: basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Clayton MJ (1997). Delphi: a technique to harness expert opinion for critical decision making tasks in education. *Educational Psychology, 17*(4), 373-386.

Cohen JJ (2007). Viewpoint: linking professionalism to humanism: what it means, why it matters. *Academic Medicine, 82*(11), 1029-1032.

Cohen JS & Patten S (2005). Well-being in residency training: a survey examining resident physician satisfaction both within and outside of residency training and mental health in Alberta. *BMC Medical Education, 5*(21).

Cohen L & Manion L (1994). *Research Methods in Education* New York: Routledge.

Cohn F & Lie D (2002). Mediating the gap between the white coat ceremony and the ethics and professionalism curriculum. *Academic Medicine, 77*(11), 1168.

Cook C, Heath F & Thompson RL (2000). A meta-analysis of response rates in Web- or internet-based surveys. *Educational and Psychological Measurement, 60*(6), 821-836.

Cook J, Dickinson H & Eccles M (2009). Response rate in postal surveys of healthcare professionals between 1996 and 2005: An observational study. *BMC Health Services Research, 9*(160).

Cornwell JM & Dunlap WP (1994). On the questionable soundness of factoring ipsative data - a response. *Journal of Occupational and Organizational Psychology, 67*, 89-100.

Cotterill S, Diaz S, Cather A & Shumway J (2006). Assessing medical student professionalism: an analysis of a peer assessment. *Medical Education Online, 11*(8).

Couper MP, Tourangeau R, Conrad F & Singer E (2006). Evaluating the effectiveness of visual analog scale: a web experiment. *Social Science Computer Review, 24*(2), 227 - 244.

Cowley J & Youngblood H (2009). *Subjective response differences between visual analogue, ordinal and hybrid response scales.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting. San Antonio.

Cronbach L & Meehl P (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281 - 302.

Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 298 - 334.

Cronbach LJ, Nageswari R & Gleser GC (1963). Theory of Generalizability: a liberation of reliability theory. *The British Journal of Statistical Psychology, 16*, 137-163.

Cruess RL & Cruess SR (2000). Medical professionalism in society. *New England Journal of Medicine, 342*(17), 1288-1289.

Cruess RL & Cruess SR (2006). Teaching professionalism: general principles. *Medical Teacher, 28*(3), 205-208.

Cruess RL & Cruess SR (2011). Commentary: professionalism, unionization, and physicians' strikes. *Academic Medicine, 86*(5), 548-551.

Cruess RL, Cruess SR & Johnston SE (2000a). Professionalism and medicine's social contract. *Journal of Bone and Joint Surgery - American Edition, 82A*(8), 1189-1194.

Cruess RL, Cruess SR & Johnston SE (2000b). Professionalism: an idea to be sustained. *Lancet, 365*(9224), 156 - 159.

Cruess RL, McIlroy JH, Ginsburg S & Steinert Y (2006). The professionalism mini-evaluation exercise: a preliminary investigation *Academic Medicine, 81*(10), s74 - 78.

Cruess SR & Cruess RL (1997). Professionalism must be taught. *British Medical Journal, 315*, 1674 - 1677.

Cruess SR, Cruess RL & Steinert Y (2008). Teaching rounds - role modelling - making the most of a powerful teaching strategy. *British Medical Journal, 336*(7646), 718-721.

Cruess SR, Cruess RL & Steinert Y (2010). Teaching professionalism across cultural and national borders: Lessons learned from an AMEE workshop. *Medical Teacher, 32*(5), 371-374.

Cruess SR, Johnston S & Cruess RL (2004). "Profession": a working definition for medical educators. *Teaching and Learning in Medicine, 16*(1), 74-76.

Cumming RG (1990). Is probability-sampling always better - a comparison of results from a quota and a probability sample survey. *Community Health Studies, 14*(2), 132-137.

Davis MH, Ben-David MF, Harden RM, Ker PHJ, McGhee C, Pippard MJ *et al.* (2001). Portfolio assessment in medical students' final examinations. *Medical Teacher, 23*(4), 357-366.

Davis RL, Wiggins MN, Mercado CC & O'Sullivan PS (2007). Defining the core competency of professionalism based on the patient's perception. *Clinical & Experimental Ophthalmology, 35*(1), 51-54.

de Leeuw ED (2001). Reducing missing data in surveys: an overview of methods. *Quality & Quantity, 35*(2), 147-160.

De Villiers MR, De Villiers PJT & Kent AP (2005). The Delphi technique in health sciences education research. *Medical Teacher, 27*(7), 639-643.

DeCoster J (1998). Overview of factor analysis. Retrieved 10 January 2011, 2011, from www.stat-help.com/notes.html.

Deehan A, Templeton L, Taylor C, Drummond C & Strang J (1997). The effect of cash and other financial inducements on the response rate of general practitioners in a national postal study. *British Journal of General Practice, 47*(415), 87-90.

Delnevo CD, Abatemarco DJ & Steinberg MB (2004). Physician response rates to a mail survey by specialty and timing of incentive. *American Journal of Preventive Medicine, 26*(3), 234-236.

Deloney LA (2003). The ICM pin: a symbol of professional development during medical education. *Teaching and Learning in Medicine, 15*(3), 215-216.

Deshpande SP (2009). A study of ethical decision making by physicians and nurses in hospitals. *Journal of Business Ethics, 90*(3), 387-397.

Deutskens E, De Ruyter K, Wetzels M & Oosterveld P (2004). Response rate and response quality of Internet-based surveys: an experimental study. *Marketing Letters, 15*(1), 21-36.

DeVellis RF (2003). *Scale Development: Thoery and Application* (Vol. 26). London SAGE Publications.

DiLullo C, Morris HJ & Kriebel RM (2009). Clinical competencies and the basic sciences: an online case tutorial paradigm for delivery of integrated clinical and basic science content. *Anatomical Sciences Education, 2*(5), 238-243.

Driessen E, van der Vleuten C, Schuwirth L, van Tartwijk J & Vermunt J (2005). The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Medical Education, 39*(2), 214-220.

Duff P (2004). Teaching and assessing professionalism in medicine. *Obstetrics & Gynecology, 104*(6), 1362-1366.

Dundee Medical School (2008). Dundee MBChB programme. Retrieved 20 January, 2012, from http://medicine.dundee.ac.uk/.

Dziuban CD & Shirkey EC (1974). When is a corralational matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin, 81*(6), 358-361.

Elit L, Hunt M, Redwood-Campbell L, Ranford J, Adelson N & Schwartz L (2011). Ethical issues encountered by medical students during international health electives. *Medical Education, 45*(7), 704-711.

Elliott DD, May W, Schaff PB, Nyquist JG, Trial J, Reilly JM *et al.* (2009). Shaping professionalism in pre-clinical medical students: Professionalism and the practice of medicine. *Medical Teacher, 31*(7), e295-302.

Epstein RM (2007). Assessment in medical education. *New England Journal of Medicine, 356*(4), 387-396.

Epstein RM & Hundert EM (2002). Defining and assessing professional competence. *Jama-Journal of the American Medical Association, 287*(2), 226-235.

Eun-ok IM & Wonshik C (2011). quota sampling in internet research: Parctical issues *Computers, Informatics, Nursing, 29*(7), 381-385.

Eva KW & Regehr G (2005). Self-assessment in the health professions: a reformulation and research agenda. *Academic Medicine, 80*(10 Suppl), S46-54.

Fard NN, Asghari F & Mirzazadeh A (2010). Ethical issues confronted by medical students during clinical rotations. *Medical Education, 44*(7), 723-730.

Farnan JM, Paro JA, Higa J, Edelson J & Arora VM (2008). The YouTube generation: implications for medical professionalism. *Perspectives in Biology and Medicine, 51*(4), 517-524.

Feudtner C, Christakis D & Christakis N (1994). Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Academic Medicine, 69*(8), 670-679.

Field A (2009). *Discovering Statistics Using SPSS* (3rd ed.). Los Angeles: SAGE Publications.

Finn G, Sawdon M, Clipsham L & McLachlan J (2009). Peer estimation of lack of professionalism correlates with low Conscientiousness Index scores. *Medical Education, 43*(10), 960-967.

Fins JJ, Gentilesco BJ, Carver A, Lister P, Acres CA, Payne R *et al.* (2003). Reflective practice and palliative care education: a clerkship responds to the informal and hidden curricula. *Academic Medicine, 78*(3), 307-312.

Francis J, Eccles M, Johnston M, Walker A, Grimshaw J, Foy R *et al.* (2004). *Constructing questionnaires based on the theory of planned behaviour: A manual for healthcare services researchers* Newcastle upon Tyne: Centre for Health Services Research, University of Newcastle.

Frank JR & Danoff D (2007). The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Medical Teacher, 29*(7), 642-647.

Fresa-Dillon KL, Cuzzolino RG & Veit KJ (2004). Professionalism: orientation exercises for incoming osteopathic medical students and developing class vision statements. *The Journal of the American Osteopathic Association, 104*(6), 251-259.

Frohna A & Stern D (2005). The nature of qualitative comments in evaluating professionalism. *Medical Education, 39*(8), 763-768.

Fryer-Edwards K, Pinsky LE & Robins L (2006). The use of portfolio to assess professionalism In Stern A (Ed.), *Measuring Medical Professionalism* (pp. 213 - 233). New York: Oxford University Press.

Gaessaroli M (2006). *Generalisabilty Theory* Paper presented at the International Summer Institute for Medical Education Research Centre for Medical Education, Dundee.

Garner J & O'Sullivan H (2010). Facebook and the professional behaviours of undergraduate medical students. *Clinical Teacher, 7*(2), 112-115.

Geer JG (1991). Do open-ended questions measure salient issues. *Public Opinion Quarterly, 55*(3), 360-370.

Gillon R (1994). Medical ethics: four principles plus attention to scope. [Review]. *British Medical Journal, 309*(6948), 184-188.

Ginsburg S & Lingard L (2011). 'Is that normal?' Pre-clerkship students' approaches to professional dilemmas. *Medical Education, 45*(4), 362-371.

Ginsburg S, Regehr G, Hatala R, McNaughton N, Frohna A, Hodges B *et al.* (2000). Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Academic Medicine, 75*(10 Suppl), S6-S11.

Ginsburg S, Regehr G & Lingard L (2003). The disavowed curriculum: understanding student's reasoning in professionally challenging situations. *Journal of General Internal Medicine, 18*(12), 1015-1022.

Ginsburg S, Regehr G & Lingard L (2004). Basing the evaluation of professionalism on observable behaviors: A cautionary tale. *Academic Medicine, 79*(10), S1-S4.

Ginsburg S, Regehr G & Mylopoulos M (2009). From behaviours to attributions: further concerns regarding the evaluation of professionalism. *Medical Education, 43*(5), 414-425.

Ginsburg SR, Regehr G & Mylopoulos M (2007). Reasoning when it counts: students' rationales for action on a professionalism exam. *Academic Medicine, 82*(10 Suppl), S40-43.

Glicken AD & Merenstein GB (2007). Addressing the hidden curriculum: understanding educator professionalism. *Medical Teacher, 29*(1), 54-57.

GMC (1993). *Tomorrow's Doctors.* London: General Medical Council. 15.

GMC (2003). *Tomorrow's Doctors* London: General Medical Council. 25 - 27.

GMC (2006). *Good Medical Practice*. London: General Medical Council. 2-34.

GMC (2009). *Tomorrow's Doctors*. London General Medical Council 25 - 29.

Goldberg JL (2008). Humanism or professionalism? The White Coat Ceremony and medical education. *Academic Medicine, 83*(8), 715-722.

Goldie J (2008). Integrating professionalism teaching into undergraduate medical education in the UK setting. *Medical Teacher, 30*(5), 513-527.

Goldie J, Dowie A, Cotton P & Morrison J (2007). Teaching professionalism in the early years of a medical curriculum: a qualitative study. *Medical Education, 41*(6), 610-617.

Goldie J, Schwartz L, McConnachie A & Morrison J (2004). The impact of a modern medical curriculum on students' proposed behaviour on meeting ethical dilemmas. *Medical Education, 38*(9), 942-949.

Gooden BR, Smith MJ, Tattersall SJN & Stockler MR (2001). Hospitalised patients' views on doctors and white coats. *Medical Journal of Australia, 175*(4), 219-222.

Gordon J (2003). Fostering students' personal and professional development in medicine: a new framework for PPD. *Medical Education, 37*(4), 341-349.

Gosling SD, Vazire S, Srivastava S & John OP (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59*(2), 93-104.

Grava-Gubins I & Scott S (2008). Effects of various methodologic strategies Survey response rates among Canadian physicians and physicians-in-training. *Canadian Family Physician, 54*(10), 1424-1430.

Graves DL, Carolyn KS & Arnold L (2002). The role of spirituality in patient care: Incorperating spirituality training into medical school curriculum *Academic Medicine, 77*(11), 1167.

Green M, Zick A & Makoul G (2009). Defining professionalism from the perspective of patients, physicians, and nurses. *Academic Medicine, 84*(5), 566-573.

Grodzinsky F (1999). The practitioner from within: revisiting the virtues. *Computers and Society, 29*(1), 9 -15.

Hafferty F (2006). Definitions of professionalism - a search for meaning and identity. *Clinical Orthopaedics and Related Research*(449), 193-204.

Hafferty FW (1998). Beyond curriculum reform: confronting medicine's hidden curriculum. *Academic Medicine, 73(4)*, 403 - 407.

Hafferty FW & Castellani B (2009). A sociological framing of medicine's modern day professional movement. *Medical Education*(43), 826 - 828.

Hafferty FW & Castlellani B (2009). A sociological framing of medicine' modern-day professionalism movement. *Medical Education, 43*, 826 - 828.

Hammond SM, O'Rourke M, Kelly M, Bennett D & O'Flynn S (2012). A psychometric appraisal of the DREEM. *BMC Medical Education, 12*(2).

Harden RM (1999). AMEE Guide No. 14: outcome-based education: Part 1- an introduction to outcome-based education. *Medical Teacher, 21*(1), 7-14

Hargittai E (2009). An update on survey measures of web-oriented digital literacy. *Social Science Computer Review, 27*(1), 130-137.

Hawkins RE, Katsufrakis PJ, Holtman MC & Clauser BE (2009). Assessment of medical professionalism: who, what, when, where, how, and ... why? *Medical Teacher, 31*(4), 348-361.

Hayens S, Richard D & Kubany E (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238 - 247.

Hilton SR & Slotnick HB (2005). Proto-professionalism: how professionalisation occurs across the continuum of medical education. *Medical Education, 39*(1), 58-65.

Hodges BD, Ginsburg S, Cruess R, Cruess S, Delport R, Hafferty F *et al.* (2011). Assessment of professionalism: recommendations from the Ottawa 2010 Conference. *Medical Teacher, 33*(5), 354-363.

Hodges D, McLachlan JC & Finn GM (2009). Exploring reflective 'critical incident' documentation of professionalism lapses in a medical undergraduate setting. *BMC Medical Education, 9*(44).

Hoff TJ (2000). Medical professionalism in society. *New England Journal of Medicine, 342*(17), 1289-1289.

Hojat M, Gonnella JS, Mangione S, Nasca TJ & Magee M (2003a). Physician empathy in medical education and practice: experience with the Jefferson scale of physician empathy. *Seminars in Integrative Medicine, 1*(1), 25 - 41.

Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M & Magee M (2002). Physician empathy: definition, components, measurement, and relationship to gender and specialty. *American Journal of Psychiatry, 159*(9), 1563-1569.

Hojat M, Mangione S, Nasca TJ, Cohen MJM, Gonnella JS, Erdmann JB *et al.* (2001). The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educational and Psychological Measurement, 61*(2), 349-365.

Hojat M, Nasca TJ, Erdmann JB, Frisby AJ, Veloski JJ & Gonnella JS (2003b). An operational measure of physician lifelong learning: its development, components and preliminary psychometric data. *Medical Teacher, 25*(4), 433-437.

Hojat M, Veloski J, Nasca TJ, Erdmann JB & Gonnella JS (2006). Assessing physicians' orientation toward lifelong learning. *Journal of General Internal Medicine, 21*(9), 931-936.

Hojat M, Veloski JJ & Gonnella JS (2009a). Measurement and correlates of physicians' lifelong learning. *Academic Medicine, 84*(8), 1066-1074.

Hojat M, Vergare MJ & Maxwell K (2009b). The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Academic Medicine, 84*(11), 1616-1616.

Hornby AS (1995) Oxford Advanced Learner's Dictionary of Current English In: Crowthier J. (ed.) 5 ed. Oxford: Oxford University Press.

Horton R (2005). Medicine: the prosperity of virtue. *Lancet, 366*(9502), 1985-1987.

Howe A, Barrett A & Leinster S (2009). How medical students demonstrate their professionalism when reflecting on experience. *Medical Education, 43*(10), 942-951.

Huddle TS (2005). Teaching professionalism: Is medical morality a competency? *Academic Medicine, 80*(10), 885 - 891.

Hudelson P (2004). Culture and quality: an anthropological perspective. *International Journal for Quality in Health Care, 16*(5), 345 - 346.

Huggett KN, Warrier R & Maio A (2008). Early learner perceptions of the attributes of effective preceptors. *Advances in Health Sciences Education : Theory and Practice, 13*(5), 649-658.

Hurwitz B & Richardson R (1997). Swearing to care: the resurgence in medical oaths. *British Medical Journal, 315*(7123), 1671-1674.

Hutchings HA & Rapport F (2012). Patient-centered professionalism. *Patient Intelligence, 4*, 11-21.

Hutchinson L (1999). Evaluating and researching the effectiveness of educational interventions. *British Medical Journal, 318*(7193), 1267-1269.

Irvine D (2001). Doctors in the UK: their new professionalism and its regulatory framework. *Lancet, 358*(9295), 1807-1810.

Irvine D (2005). Patients, professionalism, and revalidation. *British Medical Journal, 330*(7502), 1265-1268.

Jamieson S (2004). Likert scales: how to (ab)use them. *Medical Education, 38*(12), 1217-1278.

Jethwa S, Bryant P, Singh S, Jones M, Berlin A & Rosenthal J (2009). Your life in their pocket: students' behaviors regarding confidential patient information. *Family Medicine, 41*(5), 327-331.

Jha V, Bekker HL, Duffy SR & Roberts TE (2007). A systematic review of studies assessing and facilitating attitudes towards professionalism in medicine. *Medical Education, 41*(8), 822-829.

Jha V, Bekker HL, Duffy SRG & Roberts TE (2006). Perceptions of professionalism in medicine: a qualitative study. *Medical Education, 40*(10), 1027-1036.

Johnston P & Wilkinson K (2009). Enhancing validity of critical task selected for college and university programme portfolios. *National Forum of Teacher Education Journal, 19*(November 3), 1 - 6.

Joinson AN, Woodley A & Reips UD (2007). Personalization, authentication and self-disclosure in self-administered Internet surveys. *Computers in Human Behavior, 23*(1), 275-285.

Jones J & Hunter D (1995). Qualitative research: consensus methods for medical and health services research. *British Medical Journal, 311 (7001)*, 376.

Jones WS, Hanson JL & Longacre JL (2004). An intentional modeling process to teach professional behavior: students' clinical observations of preceptors. *Teaching and Learning in Medicine, 16*(3), 264-269.

Jotkowitz AB, Glick S & Porath A (2004). A physician charter on medical professionalism: a challenge for medical education. *European Journal of Internal Medicine, 15*(1), 5-9.

Kalet A, Krackov S & Rey M (2002). Mentoring for a new era. *Academic Medicine, 77*(11), 1171-1172.

Kal*et al* & Steven H (2004). Measuring professionalism attitudes and beliefs *Journal of General Internal Medicine, 19*(s1), 177.

Kao A (2006). Ethics, law and professionalism: what physicians need to know? In Stern DT (Ed.), *Measuring Medical Professionalism*. New York: Oxford University Press.

Karnieli-Miller O, Vu TR, Holtman MC, Clyman SG & Inui TS (2010). Medical students' professionalism narratives: a window on the informal and hidden curriculum. *Academic Medicine, 85*(1), 124-133.

Kellerman SE & Herold J (2001). Physician response to surveys: A review of the literature. *American Journal of Preventive Medicine, 20*(1), 61 - 67.

Kenny N, Mann K & MacLeod H (2003). Role modeling in physicians' professional formation: reconsidering an essential but untapped educational strategy. *Academic Medicine, 78*(12), 1203-1210.

Ker JS, Hesketh EA, Anderson F & Johnston DA (2006). Can a ward simulation exercise achieve the realism that reflects the complexity of everyday practice junior doctors encounter? *Medical Teacher, 28*(4), 330-334.

Kianifard F (1994). Evaluation of clinimetric scales: basic principles and methods *Journal of the Royal Statistical Society. Series D (The Statistician), 43*(4), 475-482.

Kline P (1993). *The Handbook of Psychological Testing* London: Routledge.

Knights JA & Kennedy BJ (2006). Medical school selection: screening for dysfunctional tendencies. *Medical Education, 40*(11), 1058-1064.

Kohlberg L (1973). Stages and aging in moral development - some speculations. *The Gerontologist, Winter*, 497 - 502.

Kostas TR, Jones DB, Schiefer TK, Geske JB, Carmichael SW & Pawlina W (2007). The use of a video interview to enhance gross anatomy students' understanding of professionalism. *Medical Teacher, 29*(2-3), 264-266.

Krych EH & Vande Voort JL (2006). Medical students speak: a two-voice comment on learning professionalism in medicine. *Clinical Anatomy, 19*(5), 415-418.

Kumagai AK & Lypson ML (2009). Beyond cultural competence: critical consciousness, social justice, and multicultural education. *Academic Medicine, 84*(6), 782-787.

Lambe P & Bristow D (2010). What are the most important non-academic attributes of good doctors? A Delphi survey of clinicians. *Medical Teacher, 32*(8), e347-354.

Lawshe C (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563 - 575.

Lempp H & Seale C (2004). The hidden curriculum in undergraduate medical education: qualitative study of medical students' perceptions of teaching. *British Medical Journal, 329*(7469), 770-773.

Lie D, Rucker L & Cohn F (2002). Using literature as the framework for a new course. *Academic Medicine, 77*(11), 1170.

Loh KY & Nalliah S (2010). Learning professionalism by role-modelling. *Medical Education, 44*(11), 1123.

Lozano LM, García-Cueto E & Muñiz J (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology, 4*(2), 73-79.

Lynch DC, Surdyk PM & Eiser AR (2004). Assessing professionalism: a review of the literature. *Medical Teacher, 26*(4), 366-373.

Lynn MR (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-385.

Madill A & Latchford G (2005). Identity change and the human dissection experience over the first year of medical training. *Social Science & Medicine, 60*(7), 1637-1647.

Makaul G & Altman M (2002). Early assesment of medical students' clinical skills. *Academic Medicine, 77*(11), 1156.

Marshall MN (1999). How well do GPs and hospital consultants work together? A survey of the professional relationship. *Family Practice, 16*(1), 33-38.

Martimianakis MA, Maniate JM & Hodges BD (2009). Sociological interpretations of professionalism. *Medical Education, 43*(9), 829-837.

Maudsley G, Williams EMI & Taylor DCM (2007). Junior medical students' notions of a 'good doctor' and related expectations: a mixed methods study. *Medical Education, 41*(5), 476-486.

Mazor KM, Canavan C, Farrell M, Margolis MJ & Clauser BE (2008). Collecting validity evidence for an assessment of professionalism: findings from think-aloud Interviews. *Academic Medicine, 83*(10), s9-12.

McAleer J & Chandratilake M (In press). Choosing assessment instruments. In Walsh K (Ed.), *Oxford Textbook of Medical Education* London: BMJ.

McLachlan J (2010). Measuring conscientiousness and professionalism in undergraduate medical students. *Clinical Teacher, 7*(1), 37-40.

McMillan B & Conner M (2003). Using the theory of planned behaviour to understand alcohol and tobacco use in students. *Psychology, Health and Medicine, 8*(3), 317 - 328.

McMillan B & Conner M (2009). Using the theory of planned behaviour to understand alcohol and tobacco use in students. *Psychology, Health and Medicine, 8*(3), 317 - 328.

Meade AW & Craig SB (2012). Identifying careless responses in survey data. *Psychological Methods, Advance online publication. doi: 10.1037/a0028085*.

Medical School Council (2007). *Medical Students: Professional Behaviour and Fitness to Practice*. London: General Medical Council.

Miles S & Leinster SJ (2010). Identifying professional characteristics of the ideal medical doctor: The laddering technique. *Medical Teacher, 32*(2), 136-140.

Mitra A, Jain-shukla P, Robbins A, Champion H & Durant R (2008). Difference in rate of response to web-based survey among college students. *International Journal on E-Learning, 7*(2), 265 - 281.

Monrouxe L & Rees C (2011). “It’s just a clash of cultures”: emotional talk within medical students’ narratives of professionalism dilemmas. *Advances in Health Sciences Education* (Epub ahead of print).

Monrouxe LV (2010). Identity, identification and medical education: why should we care? *Medical Education, 44*(40-49).

Monrouxe LV, Rees CE & Bradley P (2009). The construction of patients' involvement in hospital bedside teaching encounters. *Qualitative Health Research, 19*(7), 918-930.

Monrouxe LV, Rees CE & Hu W (2011). Differences in medical students' explicit discourses of professionalism: acting, representing, becoming. *Medical Education, 45*(6), 585-602.

Mooney CJ, Lurie SJ, Lyness JM, Lambert DR & Guzick DS (2010). Development of an audit method to assess the prevalence of the ACGME's general competencies in an undergraduate medical education curriculum. *Teaching and Learning in Medicine, 22*(4), 257-261.

Moore L & Savage J (2002). Participant observation, informed consent and ethical approval. [Review]. *Nurse researcher, 9*(4), 58-69.

Moseley TH, Cantrell MJ & Deloney LA (2002). Clinical skills senter sttending: an innovative senior medical school elective. *Academic Medicine, 77*(11), 1176.

Moss J & Hendry G (2002). Use of electronic surveys in course evaluation. *British Journal of Educational Technology, 33*(5), 583-592.

Mueller PS (2009). Incorporating professionalism into medical education: the Mayo Clinic experience. *Keio J Medicine, 58*(3), 133-143.

Natanzon I, Ose D, Szecsenyi J, Campbell S, Roos M & Joos S (2010). Does GPs' self-perception of their professional role correspond to their social self-image? - A qualitative study from Germany. *BMC Family Practice (Online), 11*(10).

National Library of Medicine (2009, 1 June 2009). Greek medicine: Hippocratic Oath. Retrieved 11 September 09, 2009, from http://www.nlm.nih.gov/.

Nofziger AC, Naumburg EH, Davis BJ, Mooney CJ & Epstein RM (2010). Impact of peer assessment on the professional development of medical students: a qualitative study. *Academic Medicine, 85*(1), 140-147.

Norcini JJ, Blank LL, Duffy FD & Fortna GS (2003). The Mini-CEX: a method for assessing clinical skills. *Annals of Internal Medicine, 138*(6), 476-481.

Nova Online (2009). Hippocratic Oath: modern version. Retrieved 14 September, 2009, from http://www.pbs.org/wgbh/nova/doctors/oath_modern.html.

NRS (2009). Social grade - definition and discriminatory power. Retrieved 15.09.2009, 2009, from http://www.nrs.co.uk/lifestyle.html.

O'Sullivan AJ & Toohey SM (2008). Assessment of professionalism in undergraduate medical students. *Medical Teacher, 30*(3), 280-286.

O'Sullivan H, van Mook W, Fewtrell R & Wass V (2012). Integrating professionalism into the curriculum: AMEE Guide No. 61. *Medical Teacher, 34*(2), e64-77.

O'Toole TP, Kathuria N, Mishra M & Schukart D (2005). Teaching professionalism within a community context: perspectives from a national demonstration project. *Academic Medicine, 80*(4), 339-343.

Office for National Statistics (2008). Internet access 2008: households and individuals. Retrieved 16 November, 2009, from http://www.statistics.gov.uk/.

ONS (2009). Population. Retrieved 02 September 2009, from http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Population.

Ormrod J (1999). *Human Learning* (3rd ed.). Upper Saddle River,NJ: Prentice.

Paolo AM, Bonaminio GA, Gibson C, Partridge T & Kallail K (2000). Response rate comparisons of e-mail- and mail-distributed student evaluations. *Teaching and Learning in Medicine, 12*(2), 81-84.

Papadakis AM, Loeser H & Healy K (2001). Early detection of evaluation of professionalism deficiencies in medical students: one school's approach. *Academic Medicine, 76*(11), 1100 - 1106.

Papadakis MA, Osborn EH, Cooke M & Healy K (1999). A strategy for the detection and evaluation of unprofessional behavior in medical students. University of California, San Francisco School of Medicine Clinical Clerkships Operation Committee. *Academic Medicine, 74*(9), 980-990.

Papageorgiou A, Miles S, Fromage M, Kemmy J & Leinster SJ (2011). Cross-sectional evaluation of a longitudinal consultation skills course at a new UK medical school. *BMC Medical Education, 11* (55).

Parker M (2006). Assessing professionalism: theory and practice. *Medical Teacher, 28*(5), 399-403.

Pasquale SJ & Pugnaire MP (2002). Preparing medical students to teach. *Academic Medicine, 77*(11), 1175 - 1176.

Passmore K & Leung WC (2002). Defensive practice among psychiatrists: a questionnaire survey. *Postgraduate Medical Journal, 78*(925), 671-673.

Patenaude J, Niyonsenga T & Fafard D (2003). Changes in students' moral development during medical school: a cohort study. *Canadian Medical Association Journal, 168*(7), 840-844.

Pawlina W (2006). Professionalism and anatomy: how do these two terms define our role? *Clinical Anatomy, 19*(5), 391-392.

Pell G (2005). Use and misuse of Likert scales. *Medical Education, 39*(9), 970 - 971.

Polit D & Beck C (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*(5), 489 - 497.

Porter SR, Whitcomb ME & Weitzer WH (2004). Multiple surveys of students and survey fatigue *New Directions for Institutional Research, Spring* 63 - 73.

Porter SR & Whitecomb ME (2003). The impact of contact type on web survey resposne rates. *Public Opinion Quarterly, 67* (4), 579 - 588.

Powell C (2003). The Delphi technique: myths and realities. *Journal of Advanced Nursing, 41*(4), 376-382.

Prislin MD, Lie D, Shapiro J, Boker J & Radecki S (2001). Using standardized patients to assess medical students' professionalism. *Academic Medicine, 76*(10), s90-92.

Quaintance JL, Arnold L & Thompson GS (2008). Development of an instrument to measure the climate of professionalism in a clinical teaching environment. *Academic Medicine, 83*(10 Suppl), s5-8.

Quaintance JL, Arnold L & Thompson GS (2010). What students learn about professionalism from faculty stories: an "appreciative Inquiry" approach. *Academic Medicine, 85*(1), 118-123.

Rabow MW, Wrubel J & Remen RN (2009). Promise of professionalism: personal mission statements among a national cohort of medical students. *Annals of Family Medicine, 7*(4), 336-342.

Randall DM & Gibson AM (1991). Ethical decision making in the medical profession: an application of the theory of planned behaviour. *Journal of Bussiness Ethics, 10*, 111 - 122.

Randolph J (2008). Online kappa calculator. Retrieved November 20, 2010, 2010, from http://justus.randolph.name/kappa.

Randolph JJ (2005, 14 - 15 October 2005.). *Free-marginal multirater kappa (multirater free): an alternative to Fleiss' fixed-marginal multirater kappa.* Paper presented at the University of Joensuu Learning and Instruction Symposium, Joensuu, Finland.

Ranganath KA, Smith CT & Nosek BA (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology, 44*(2), 386-396.

Rankin WL & Grube JW (1980). A comparison of ranking and rating procedures for value system measurement. *European Journal of Social Psychology, 10*(3), 233-246.

RCP (2005). *Doctors in society: medical professionalism in a changing world. Report of a Working Party of the Royal College of Physicians of London*. London: Royal College of Physicians.

RCP (2009). Multisource feedback. Retrieved 10 January, 2012, from http://old.rcplondon.ac.uk/clinical-standards/revalidation/Pages/tools.aspx.

RCPSC (2005). The CanMEDS Framework: Royal College of Physicians and Surgeons in Canada. Retrieved 10 January 2012, from http://www.royalcollege.ca/public/canmeds/framework.

Rees C & Shepherd M (2005a). The acceptability of 360-degree judgements as a method of assessing undergraduate medical students' personal and professional behaviours. *Medical Education, 39*(1), 49-57.

Rees C & Shepherd M (2005b). Students' and assessors' attitudes towards students' self-assessment of their personal and professional behaviours. *Medical Education, 39*(1), 30-39.

Rees CE & Knight LV (2007). The trouble with assessing students' professionalism: theoretical insights from sociocognitive psychology. *Academic Medicine, 82*(1), 46-50.

Rees CE & Monrouxe LV (2011). Medical students learning intimate examinations without valid consent: a multicentre study. *Medical Education, 45*(3), 261-272.

Regis T, Steiner MJ, Ford CA & Byerley JS (2011). Professionalism expectations seen through the eyes of resident physicians and patient families. *Pediatrics, 127*(2), 317-324.

Reid R, Bruce D, Allstaff K & McLernon D (2006). Validating the readiness for interprofessional learning scale (RIPLS) in the postgraduate context: are health care professionals ready for IPL? *Medical Education, 40*(5), 415-422.

Remen RN & Rabow MW (2005). The healer's art: professionalism, service and mission. *Medical Education, 39*(11), 1167-1168.

Ringsted C, Hansen TL, Davis D & Scherpbier A (2006). Are some of the challenging aspects of the CanMEDS roles valid outside Canada? *Medical Education, 40*(8), 807-815.

Riquelme A, Oporto M, Oporto J, Mendez JI, Viviani P, Salech F *et al.* (2009). Measuring students' perceptions of the educational climate of the new curriculum at the Pontificia Universidad Catolica de Chile: performance of the Spanish translation of the Dundee Ready Education Environment Measure (DREEM). *Education for Health (Abingdon), 22*(1), 112.

Robb K, Wardle J, Stubbings S, Ramirez A, Austoker J, Macleod U *et al.* (2010). Ethnic disparities in knowledge of cancer screening programmes in the UK. *Journal of Medical Screening, 17*(3), 125-131.

Roberts LW, Warner TD, Hammond KA, Geppert CM & Heinrich T (2005). Becoming a good doctor: perceived need for ethics training focused on practical and professional development topics. *Academic Psychiatry, 29*(3), 301-309.

Robins LS, Braddock CH, 3rd & Fryer-Edwards KA (2002). Using the American Board of Internal Medicine's "Elements of Professionalism" for undergraduate ethics education. *Academic Medicine, 77*(6), 523-531.

Roff S (2005). The Dundee Ready Educational Environment Measure (DREEM) - a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Medical Teacher, 27*(4), 322-325.

Roff S, Chandratilake M, McAleer S & Gibson J (2011). Preliminary benchmarking of appropriate sanctions for lapses in undergraduate professionalism in the health professions. *Medical Teacher, 33*(3), 234-238.

Roff S, Chandratilake M, McAleer S & Gibson J (2012). Medical student rankings of proposed sanction for unprofessional behaviours relating to academic integrity: results from a Scottish medical school. *Scottish Medical Journal, 57*(2), 76-79.

Roff S & Dherwani K (2011). Development of inventory for polyprofessionalism lapses at the proto-professional stage of health professions education together with recommended responses. *Medical Teacher, 33*(3), 239-243.

Roff S, McAleer S, Chandratilake M & Dherwani K (2009). Developing stage-specific, consistent, reliable and valid learning, teaching and assessment methods for PolyProfessionalism in the health professions. *01, summer*, 18 - 19.

Roff S, McAleer S, Harden RM, Al-Qahtani M, Ahmed AU, Deza H *et al.* (1997). Development and validation of the Dundee Ready Education Environment Measure (DREEM). *Medical Teacher, 19*(4), 295-299.

Rogers W & Ballantyne A (2010). Towards a practical definition of professional behaviour. *Journal of Medical Ethics, 36*(4), 250-254.

Rose GL, Rukstalis MR & Schuckit MA (2005). Informal mentoring between faculty and medical students. *Academic Medicine, 80*(4), 344-348.

Rosenthal S, Howard B, Schlussel YR, Herrigel D, Smolarz BG, Gable B *et al.* (2011). Humanism at heart: preserving empathy in third-year medical students. *Academic Medicine, 86*(3), 350-358.

Rowe G & Wright G (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting, 15*(4), 353-375.

Rubio DM, Berg-Weger M, Tebb SS, Lee ES & Rauch S (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-104.

Russell JA (1979). Affective space is bipolar. *Journal of Personality and Social Psychology, 37*(3), 345-356.

Russell PC (2002). The White Coat Ceremony: turning trust into entitlement. *Teaching and Learning in Medicine, 14*(1), 56-59.

Satterwhite WM, 3rd, Satterwhite RC & Enarson CE (1998). Medical students' perceptions of unethical conduct at one medical school. *Academic Medicine, 73*(5), 529-531.

Sax LJ, Gilmartin SK & Bryant AN (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409-432.

Scavenius M, Schmidt S & Klazinga N (2006). Genesis of the professional-patient relationship in early practical experience: qualitative and quantitative study. *Medical Education, 40*(10), 1037-1044.

Schein EH (1984). Coming to a new awareness of organisational culture. *Sloan Management Review, 25*(2), 3 - 15.

Schuman H & Presser S (1996). *Questionas and Answers in Attitude Surveys.* London: SAGE publications.

Scott T, Mannion R, Davies H & Marshall M (2003a). The quantitative measurement of organisational culture in healthcare: a review of the available instruments *Health Services Research, 38*(3), 923 - 945.

Scott T, Mannion R, Marshall M & Davies H (2003b). Does organisational culture influence health care performance? a review of the evidence. *Journal of Health Services Research and Policy, 8*(2), 105 - 117.

Seago BL, Schlesinger JB & Hampton CL (2002). Using a decade of data on medical student computer literacy for strategic planning. *Journal of the Medical Library Association, 90*(2), 202-209.

Sehiralti M, Akpinar A & Ersoy N (2010). Attributes of a good physician: what are the opinions of first-year medical students? *Journal of Medical Ethics, 36*(2), 121-125.

Seymour RA, Simpson JM, Charlton JE & Phillips ME (1985). An evaluation of length and end-phrase of visual analog Ssales in dental pain. *Pain, 21*(2), 177-185.

Shankar A, Conner M & Bodnasky HJ (2007). Can the theory of planned behaviour predicts maintenance of a frequently repeated behaviour? *Psychology, Health and Medicine, 12*(2), 213 - 224.

Shapiro J, Morrison E & Boker J (2004). Teaching empathy to first year medical students: evaluation of an elective literature and medicine course. *Education for Health (Abingdon), 17*(1), 73-84.

Shapiro J, Nguyen VP, Mourra S, Boker JR, Ross M, Thai TM *et al.* (2009). Relationship of creative projects in anatomy to medical student professionalism, test performance and stress: an exploratory study. *BMC Medical Education, 9*(65).

Shapiro J & Rucker L (2003). Can poetry make better doctors? Teaching the humanities and arts to medical students and residents at the University of California, Irvine, College of Medicine. *Academic Medicine, 78*(10), 953-957.

Shavelson RJ, Webb NM & Rowley GL (1989). Generalizability Theory. *American Psychologist, 44*(6), 922-932.

Sheehan KB (2001). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication (Online), 6*(2).

Sheeran P, Trafimow D & Armitage CJ (2003). Predicting behaviour from perceived behaviour control: tests of the accuracy assumption of the theory of planned behaviour. *British Journal of Social Psychology, 42*(3), 393 - 410.

Sheppard BH, Hartwick J & Warshaw PR (1988). The theory of reasoned action: a meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*(3), 325 - 343.

Shih TH & Fan XT (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods, 20*(3), 249-271.

Shrank WH, Reed VA & Jernstedt GC (2004). Fostering professionalism in medical education: a call for improved assessment and meaningful incentives. *Journal of General Internal Medicine, 19*(8), 887-892.

Singer E & Couper MP (2008). Do incentives exert undue influence on survey participation? Experimental evidence. *Journal of Empirical Research on Human Research Ethics, 3*(3), 49-56.

Skochelak S, Thaler S & Gjerde C (2001). The interdisciplinary generalist curriculum project at the University of Wisconsin Medical School: the generalist partners program. *Academic Medicine, 76*(4 Suppl), s131-133.

Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ *et al.* (1990). Attitudes to chemotherapy - comparing views of patients with cancer with those of doctors, nurses, and general public. *British Medical Journal, 300*(6737), 1458-1460.

Smith TMF (1983). On the validity of inferences from non-random samples. *Journal of the Royal Statistical Society Series a-Statistics in Society, 146*(4), 394-403.

Soriano R, Rosales SP, García R, García-García E & Méndez JP (2011). High knowledge about obesity and its health risks, with the exception of cancer, among mexican individuals *Journal of Cancer Education (online), DOI: 10.1007/s13187-011-0298-8*,

Sox CH (2007). The ethical foundations of professionalism; a sociologic history. *Chest, 131*(5), 1532 - 1540.

Sox H (2002). Medical professionalism in the new millennium: a physician charter. *Annals of Internal Medicine, 136*(3), 243 - 246.

Spencer JA & Jordan RK (1999). Learner centred approaches in medical education. *British Medical Journal, 318*(7193), 1280-1283.

Speyer R, Pilz W, van Der Kruis J & Brunings JW (2011). Reliability and validity of student peer assessment in medical education: A systematic review. *Medical Teacher, 33*(11), e572-585.

Starkweather J & Herrington R (2011). Principal components analysis in SPSS. *Research and Statistical Support.* Retrieved 02 May 2012, 2012, from http://www.unt.edu/rss/class/Jon/SPSS_SC/Module9/M9_PCA/SPSS_M9_PCA1.htm.

Stephenson AE, Adshead LE & Higgs RH (2006). The teaching of professional attitudes within UK medical schools: reported difficulties and good practice. *Medical Education, 40*(11), 1072-1080.

Stirrat GM, Johnston C, Gillon R & Boyd K (2010). Medical ethics and law for doctors of tomorrow: the 1998 consensus statement updated. *Journal Medical Ethics, 36*(1), 55-60.

Swartz WJ (2006). Using gross anatomy to teach and assess professionalism in the first year of medical school. *Clinical Anatomy, 19*(5), 437-441.

Swick HM (2000). Toward a normative definition of medical professionalism. *Academic Medicine, 75*(6), 612-616.

Swick HM (2006). Medical professionalism and the clinical anatomist. *Clinical Anatomy, 19*(5), 393-402.

Swick HM, Szenas P, Danoff D & Whitcomb ME (1999). Teaching professionalism in undergraduate medical education. *Journal of American Medical Association, 282*(9), 830-832.

Szauter K, Williams B, Ainsworth M, Callaway M, Bulik R & Camp M (2003). Student perceptions of professional behavior of faculty physicians. *Medical Education Online, 8*(17), 1-15.

Tavakol S, Dennick R & Tavakol M (2011a). Empathy in UK medical students: differences by gender, medical year and specialty interest. *Education for Primary Care, 22*(5), 297-303.

Tavakol S, Dennick R & Tavakol M (2011b). Psychometric properties and confirmatory factor analysis of the Jefferson Scale of Physician Empathy. *BMC Medical Education, 11*(54).

Teplitsky PE (2002). Perceptions of Canadian dental faculty and students about appropriate penalties for academic dishonesty. *Journal of Dental Education, 66*(4), 485-506.

Thrush CR, Spollen JJ, Tariq SG, Williams DK & Ii JM (2011). Evidence for validity of a survey to measure the learning environment for professionalism. *Medical Teacher, 33*(12), e683-688.

Todhunter S, Cruess SR, Cruess RL, Young M & Steinert Y (2011). Developing and piloting a form for student assessment of faculty professionalism. *Advances in Health Sciences Education, 16*(2), 223-238.

Towle A & Hoffman J (2002). An advanced communication skills course for fourth-year, post-clerkship students. *Academic Medicine, 77*(11), 1165 - 1166.

Tsai TC, Lin CH, Harasym PH & Violato C (2007). Students' perception on medical professionalism: the psychometric perspective. *Medical Teacher, 29*(2-3), 128-134.

Tsugawa Y, Ohbu S, Cruess R, Cruess S, Okubo T, Takahashi O *et al.* (2011). Introducing the Professionalism Mini-Evaluation Exercise (P-MEX) in Japan: results from a multicenter, cross-sectional Study. *Academic Medicine, 86*(8), 1026-1031.

van-de-Camp K, Vernooij-Dassen MJFJ, Grol RPTM & Bottema BJAM (2004). How to conceptualise professionalism: a qualitative study. *Medical Teacher, 26*(8), 696 - 702.

van de Camp K, Vernooij-Dassesn M, Grol R & Bottema BJAM (2006). Professionalism in general practice: development of an instrument to assess professional behaviour in general practice. *Medical Education, 40*(1), 43 - 50.

van der Vleuten C (1996). The assessment of professional competence: developments, research and practical implications. *Advances in Health Science Education, 1*(1), 41 - 67.

van Mook W, Gorter S, O'Sullivan H, Wass V, Schuwirth LW & van der Vleuten CPM (2009a). Approaches to professional behaviour assessment: Tools in the professionalism toolbox. *European Journal of Internal Medicine, 20*(8), e153-157.

van Mook W, van Luijk S, O''Sullivan H, Wass V, Schuwirth L & van der Vleuten C (2009b). General considerations regarding assessment of professional behaviour. *European Journal of Internal Medicine, 20*(4), e90 - 95.

van Mook WN, van Luijk SJ, Fey-Schoenmakers MJ, Tans G, Rethans JJ, Schuwirth LW *et al.* (2010). Combined formative and summative professional behaviour assessment approach in the bachelor phase of medical school: a Dutch perspective. *Medical Teacher, 32*(12), e517-531.

van Mook WN, van Luijk SJ, O'Sullivan H, Wass V, Harm Zwaveling J, Schuwirth LW *et al.* (2009c). The concepts of professionalism and professional behaviour: conflicts in both definition and learning outcomes. *European Journal of Internal Medicine, 20*(4), e85-89.

van Voorhis CRW & Morgan BL (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology, 3* (2), 43 -50.

van Zanten M, Boulet JR, Norcini JJ & McKinley D (2005). Using a standardised patient assessment to measure professional attributes. *Medical Education, 39*(1), 20-29.

Varga-Atkins T, Dangerfield P & Brigden D (2010). Developing professionalism through the use of wikis: a study with first-year undergraduate medical students. *Medical Teacher, 32*(10), 824-829.

Veloski JJ, Fields SK, Boex JR & Blank LL (2005). Measuring professionalism: a review of studies with instruments reported in the literature between 1982 and 2002. *Academic Medicine, 80*(4), 366-370.

Verkerk MA, De Bree MJ & Mourits MJE (2007). Reflective professionalism: Interpreting CanMEDS' "professionalism". *Journal of Medical Ethics, 33*(11), 663-666.

Wagner P, Hendrich J, Moseley G & Hudson V (2007). Defining medical professionalism: a qualitative study. *Medical Education, 41*(3), 288-294.

Wagner PJ, Lentz L & Heslop SDM (2002). Teaching communication skills: a skills-based approach. *Academic Medicine, 77*(11), 1164.

Walker A, Watson M, Grimshaw J & Bond C (2004). Applying the theory of planned behaviour to pharmacists' be;ief and intention about the treatment of vaginal candidiasis with non-prescription medicines. *Family Practice, 21*(6), 670 - 676.

Wallenstein J, Heron S, Santen S, Shayne P & Ander D (2010). A core competency-based objective structured clinical examination (OSCE) can predict future resident performance. *Academic emergency Medicine, 17 Suppl 2*, s67-71.

Wass V (2006). Doctors in society: medical professionalism in a changing world *Clinical Medicine, 6*(1), 109 - 113.

Wass V, van der Vleuten C, Shatzer J & Jones R (2001). Assessment of clinical competence. *Lancet, 357*(9260), 945-949.

Watkins MW & Pacheco M (2000). Interobserver agreement in behavioural research: Important and calculation *Journal of Behavioural Education, 10*(4), 205 - 212.

Watmough S, O'Sullivan H & Taylor D (2009). Graduates from a traditional medical curriculum evaluate the effectiveness of their medical curriculum through interviews. *BMC Medical Education, 9*(64).

Wear D & Castellani B (2000). The development of professionalism: curriculum matters. *Academic Medicine, 75*(6), 602-611.

Wetzel AP, Mazmanian PE, Hojat M, Kreutzer KO, Carrico RJ, Carr C *et al.* (2010). Measuring medical students' orientation toward lifelong learning: a psychometric evaluation. *Academic Medicine, 85*(10 Suppl), s41-44.

Whitehouse A, Hassell A, Wood L, Wall D, Walzman M & Campbell I (2005). Development and reliability testing of TAB a form for 360 degrees assessment of Senior House Officers' professional behaviour, as specified by the General Medical Council. *Medical Teacher, 27*(3), 252-258.

Wiggins MN, Coker K & Hicks EK (2009). Patient perceptions of professionalism: implications for residency education. *Medical Education, 43*(1), 28-33.

Wiggleton C, Petrusa E, Loomis K, Tarpley J, Tarpley M, O'Gorman ML *et al.* (2010). Medical students' experiences of moral distress: development of a web-based survey. *Academic Medicine, 85*(1), 111-117.

Wilkes M & Raven BH (2002). Understanding social influence in medical education. *Academic Medicine, 77*(6), 481-488.

Wilkinson TJ, Wade WB & Knock LD (2009). A blueprint to assess professionalism: results of a systematic review. *Academic Medicine, 84*(5), 551-558.

Williams SJ & Calnan M (2002). Convergence and divergence: Assessing criteria of consumer satisfaction across general practice, dental and hospital settings. *Social Science and Medicine, 33*(6), 707 - 716.

Windish DM, Paulman PM, Goroll AH & Bass EB (2004). Do clerkship directors think medical students are prepared for the clerkship years? *Academic Medicine, 79*(1), 56-61.

Wren D (1999). School culture; exploring the hidden curriculum. *Adolescence, 34*(Fall), 593 - 596.

Wynd CA, Schmidt B & Schaefer MA (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*(5), 508-518.

Zaini RG, Bin Abdulrahman KA, Al-Khotani AA, Al-Hayani AM, Al-Alwan IA & Jastaniah SD (2011). Saudi Meds: a competence specification for Saudi medical graduates. *Medical Teacher, 33*(7), 582-584.

Zilwa Dd (2007). Organisational culture and values and the adaptation of academic units in Autralian universities *Higher Education, 54*(4), 557 - 574.

Zink T, Halaas GW & Brooks KD (2009). Learning professionalism during the third year of medical school in a 9-month-clinical rotation in rural Minnesota. *Medical Teacher, 31*(11), 1001-1006.

Appendices